

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:29:58 ; Search time 51 Seconds  
(without alignments)  
1961.521 Million cell updates/sec

Title: US-10-677-980-2

Perfect score: 6481

Sequence: 1 MKGVFNIFLIPLIFLYNVI.....VOETNISDYSEYNKRMV 1210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.\*

1: /cgm2\_6/ptodata/1/iaa/5\_COMB.pep.\*

2: /cgm2\_6/ptodata/1/iaa/6\_COMB.pep.\*

3: /cgm2\_6/ptodata/1/iaa/H\_COMB.pep.\*

4: /cgm2\_6/ptodata/1/iaa/PCUTUS\_COMB.pep.\*

5: /cgm2\_6/ptodata/1/iaa/RE\_COMB.pep.\*

6: /cgm2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1557	24.0	1435	1	US-08-568-459A-4
2	1557	24.0	1435	1	US-08-487-826B-4
3	1557	24.0	1435	2	US-09-210-288-4
4	1557	24.0	1435	2	US-10-153-273-4
5	768	11.9	1115	6	5198347-6
6	761.5	11.7	749	1	US-08-568-459A-6
7	761.5	11.7	749	1	US-08-487-826B-6
8	761.5	11.7	749	2	US-09-210-288-6
9	761.5	11.7	749	2	US-10-153-273-6
10	754	11.6	1115	1	US-08-568-459A-2
11	754	11.6	1115	1	US-08-487-826B-2
12	754	11.6	1115	2	US-09-210-288-2
13	754	11.6	1115	2	US-10-153-273-2
14	655	10.1	778	6	5198347-4
15	601.5	9.3	3542	2	US-10-087-013-2
16	498.5	7.7	3060	1	US-08-487-826B-14
17	495.5	7.6	2710	1	US-08-568-459A-12
18	495.5	7.6	2710	1	US-08-487-826B-12
19	495.5	7.6	2710	2	US-09-210-288-12
20	495.5	7.6	2710	2	US-10-153-273-12
21	432	6.7	1507	2	US-08-929-329-5
22	363	5.6	407	2	US-10-087-013-8
23	360.5	5.6	921	1	US-08-568-459A-8
24	360.5	5.6	921	1	US-08-487-826B-8
25	360.5	5.6	921	2	US-09-210-288-8
26	360.5	5.6	921	2	US-10-153-273-8
27	353.5	5.5	2182	1	US-08-487-826B-16

28	348	5.4	294	2	US-10-087-013-7	Sequence 7, Appli
29	333.5	5.1	351	2	US-10-087-013-9	Sequence 9, Appli
30	327.5	5.1	308	2	US-10-087-013-11	Sequence 11, Appli
31	325	5.0	700	1	US-08-568-459A-10	Sequence 10, Appli
32	325	5.0	700	1	US-08-487-826B-10	Sequence 10, Appli
33	325	5.0	700	2	US-09-210-288-10	Sequence 10, Appli
34	325	5.0	700	2	US-10-153-273-10	Sequence 10, Appli
35	319.5	4.9	311	2	US-10-087-013-10	Sequence 10, Appli
36	309	4.8	242	2	US-08-929-329-7	Sequence 7, Appli
37	308	4.8	2184	2	US-09-417-485D-6	Sequence 6, Appli
38	301.5	4.7	277	1	US-08-568-459A-15	Sequence 15, Appli
39	301.5	4.7	277	1	US-08-487-826B-27	Sequence 27, Appli
40	301.5	4.7	277	2	US-09-210-288-15	Sequence 15, Appli
41	301.5	4.7	277	2	US-10-153-273-15	Sequence 15, Appli
42	295	4.6	2391	1	US-08-446-855A-2	Sequence 2, Appli
43	295	4.6	2391	2	US-09-150-741-2	Sequence 2, Appli
44	285.5	4.4	282	1	US-08-568-459A-16	Sequence 16, Appli
45	285.5	4.4	282	1	US-08-487-826B-28	Sequence 28, Appli

ALIGNMENTS

RESULT 1

US-08-568-459A-4

; Sequence 4, Application US/08568459A

; Patent No. 5849306

; GENERAL INFORMATION:

; APPLICANT: Sim, Kim L.

; APPLICANT: Chitnis, Chetan

; APPLICANT: Miller, Louis H.

; APPLICANT: Peterson, David S.

; APPLICANT: Su, Xin-zhaun

; APPLICANT: Wellem, Thomas E.

; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe Martens Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: California

; COUNTRY: US

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/568,459A

; FILING DATE: 07-DEC-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER: NIH121.001CP1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 235-0176

; TELEFAX: (619) 235-0176

; INFORMATION FOR SEQ ID NO: 4:

; LENGTH: 1435 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Plasmodium falciparum

; US-08-568-459A-4

Query Match

24.0%; Score 1557; DB 1; Length 1435;





```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
US-09-210-288-4

Query Match      24.0%; Score 1557; DB 2; Length 1435;
Best local Similarity 28.5%; Pred. No. 1.2e-98;
Matches 410; Conservative 214; Mismatches 467; Indels 348; Gaps 42;

QY 1 MKGYNYIYFLPLIFLVNIRINESIIGRTLYNRQDESSDISRVNSPELANNNHKNLYDS 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MKCNISYIFASFVLYFAKAREYDI-----KENEFLDYVKEKFNELDKKKYGVNQKT 55
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 DYEDVNNKLINSFVENKSVKKGRSLSPINN-----KTKSY---DIIPPSYSYRDNKFNSL 112
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 56 D-----KKIFTFIENK-----LDILANSKFNKRWKSYGTPDNIDKNNSLIN----- 96
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 113 SENEDSGNTNSNFPANTSEISIGONKQ-----YTFIQKTH----- 150
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 97 KHNEEENFNNNYOSFLSTS--SLIKONKYVPINAVRVSRILSFLDSRINNGRNTSSNNEV 154
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 151 LFACGIRKRSIKWICRENSKIT--VCVDPDKIOLCVANFLNSRLTWEKEKEIFLISVNT 209
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 155 LSNCREKRGKMKDCKKNDORSNTVCIPDRRIQICIVNLISIIKTYRKETWKDHFIEASKK 214
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 210 EAKLLYNNKNGKQPSIFCNELRNSFSDFRSFIDGDMDFGNTDRVKGYINTKFSDYKE 269
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 215 ESQLLAKNDKNYNSKFCNDLKNISFLDYGHLAGNDMDFGYSYTKAENKIQEVFKGAHGE 274
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 270 KNVEKLANIKKEWWEKANKANLNNIIVNHKNTISKECAIIPASEPQINLIKWNENFLM 329
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 275 ISEHKIKNFRKWNWNEFREKLWEAMLSSEKNNIN--NCKNIPQEBELQITQWIKEMHGESFL 333
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 330 EKKELFLNIDKCVENKYYACFGGCRLPCCSYTSFMKSKTQMEVLNLYKKNSGVDK 389
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 334 ERNRSKLPSKCKQNTLYEACEKECIDPCMVKYRDWIIIRSEFWHTLSKEYETQK--VPK 391
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 390 NNFLLDLFK--KNNKND--LDDFKKN--EKEYDDLCDCRYTATIIKSFNLGPAKNDVDIASQ 445
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 392 ENAENYLIKISENKDAAKVSILLNCAEYSKYCDCKHTTLVKSVLNGNDTIKEKREH 451
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 446 INVNDLPGGPNYKS--NNEKSNCTGTFTNKPFGTCPTCEPRRQPLCLGRTYLLHRGHEED 503
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 452 IDLDPSKFGCDKNSVDNTKVMCKNPKYILSKDVCVPPRRQBLCLGNIDRIYDKNLLM 511
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 504 YKEHLGASIEAQLAKVYKEDENALCSIIQNSYADLADIIKSGDIIKDYKQKMEEN 563
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 512 IKEHILAIAIYESRILKRYKNDKDEKVKLIINTKFPADIRDIIGGTDYWDNLSRKLVGK 571
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 564 LNKVNDKGRNEESLKIIFREKWDENKENVKVMNSAVLNKNETCKOYDKFQKIPQFLRWF 623
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 572 INTNSKYVHRKNDKLFREDEWVKVIKDVVNVISWFKDKTVCKE--DDIENIFQFPRWF 630
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 624 KEGDDPCERKEKIEYSFESFKVECKKCCODCENTCKNCKSEYKWKIDLKSEYKQV--- 680
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 631 SEWGGDYCODKTKMI---ETLKVCEKPKCEDDNCCKSKCNSEYKWKSEYKQV--- 687
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 681 DKYTKDNKKMYDNI DSVQKNEANVLYKESKECKOVNDDKIPNESPNEDMCKKDE 740
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 688 QEYQKGNKYMY-----SBFKSIKPEVILKYSEKCSNLFNFEDEFKBEHLSDYKKNCTWCPE 744
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 741 IKYL----- 746
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 745 VKDVPISIRNEQTSQEAPEENTEIAHRTETPSISEGPKGNEQKERDDDSLKLSVSP 804
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 747 -----IKYPKT----- 752
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 805 ENSRPETDAKDTSNLLKLKGDVDISMPKAVIGSSPNDNINVTEQGDNISGVNSKPLSDDV 864
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 753 ---KHDLYDID-----TFSDTFGDGTPIS----- 774
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```













Db 485 DQWTRKKNQWDLNPKFISVKNAB-----KVQTAGIVTPYDILKQELDEFNEVAFE 536  
Qy 721 DKIFNESYEVEDWCKCKDEKYLNEIKYPTKTHDIYDIDTFSTFG-DGTPI----- 772  
Db 537 NEI-NKRDGAYIELC-----VCSVEEAK-KNTQEVVTVNVDNAAKSQATNSNPIQPDVSS 589  
Qy 773 -----SINANINEQSGKDTNTGNSETSD-----SPVSHE-----PESDAAINV 812  
Db 590 KAEKVPDGSHTGNV---SGQDSSTTKAVTGDQNGNQTPAESDVQESDIAESVSAKV 646  
Qy 813 E-----KLSGD-----ESSSETRGILINDPVSNTNNVNEVHDASNTQGSV----- 852  
Db 647 DPQKSVSKRSDTASVTGIAEAGENLGASNSRPSSESTVEANSPCDDTVNSASIPVVSGE 706  
Qy 853 -----SNTSDITNGHSSSLNRTNAQDIKIGRSS-----NEQSDNQENSSHSS 896  
Db 707 NPLVTPYNGLRHSKONSDSGPAESMANPDSNSK-----GETGKGQDNMAKATKSSNSS 762  
Qy 897 DNGSLT-----IGQVPSEDNTQNTYDSQNPHRDTPNALASLPDSDKINEIEG 944  
Db 763 DGTSSATGDTTDAVDREINKGVPEDRDKTVGSKDGGEDNSANKDAATVUGEDIRENSA 822  
Qy 945 FDSRD-----SENG-----RGDTTN---THDVRRTNIV 971  
Db 823 GGSTNDRSKNDTEKNGASTPDSKQSEDATLSKTESLESTESGDRTTNDTNSLENKNG 882  
Qy 972 SERVNSHDFIRNGMANNNAHOVIT-----QIENNGIIRGOESAGNSVNYKDNPK 1023  
Db 883 KBDLQKHDFKSNTPNEBPSDQTTDAEGHRRDSIKND---KAERRKHMKNKOTFTKNTN 939  
Qy 1024 RSNFSSN---DHKNQIOEYNSRDTKRVREIILKLSKONKNENYSMEYCTYSBERNSSP 1080  
Db 940 SHLNSNNLSNGKLDIKEYKRVKATREDIILMSSVRKNNISLEYCNSVEDKISS- 998  
Qy 1081 GPCSRERKKLCCQISDYCLFYNFYSLEYNYCNKSEIKSPYKCFKSEG-QSSIPYFAA 1139  
Db 999 NTCRSREKSNLCCSDPCLNYFDVYSYELSKMGKGFEDPSYKFTKGKPKDKTYFAA 1058  
Qy 1140 GGILVIVILLSSASRMKSNHEEDIGESNTEATFER-NYLNKLSRI-----FNQEOVE 1193  
Db 1059 AGAL--LILLIASRKMKNDS-----EATFNEFEYCDNIHRIPLMPNNIEHMOP 1108  
Qy 1194 TNISDYS 1200  
Db 1109 STPLDYS 1115

## RESULT 12

US-09-210-288-2  
; Sequence 2, Application US/09210288  
; Patent No. 6392026  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellens, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/210,288  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fuller, Michael  
; REGISTRATION NUMBER: 36,516  
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1115 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Plasmodium vivax  
; US-09-210-288-2

Query Match 11.6%; Score 754; DB 2; Length 1115;

Best Local Similarity 24.5%; Pred. No. 2.5e-43;

Matches 296; Conservative 177; Mismatches 416; Indels 318; Gaps 57;

Qy 199 KPEILFISVTEAKLLVKNKEGKDPISFCNELRNSPDSFRSFFGDDMDFGGNTDRVKGY 258  
Db 22 YKDFISIT-----LINVHEGKYLILKR-----KLEKANNRDCVNCFF 59  
Qy 259 IN-TKFSDDYKKEKVEKLNKIKKEWE-KKANLNWNIWVHKGNIKECAIIPAEEPQI 316  
Db 60 LHSQVNNVLLERTITLLECKNEYVKGNYKL-----AKGH-----HC-----VEEDNL 105  
Qy 317 NLMIKWNENFLMEKKELFLNIDKCKVENKRYEACFGGCLRPCS-SYTSFMKSKTOMEV 375  
Db 106 ERMLOGTNER-----RSEENIKYK--YGVTELKIKYQAWNGKRSRILKES 149  
Qy 376 L-----TNLYKKNSGVDKNNFLNDFPKQNNKNDLDDFFKNEKEYVDLDCDXY----- 423  
Db 150 IYGAHNFPGNSYMEKGOGDKTG-----BEKQGEHKTDKTDNGKGANLVMLDYETSSN 204  
Qy 424 -----TATIKSFLNGPAKN-----DVD-----IASQINVNDLRGF-----G 455  
Db 205 GQPAGTLDNVLEFVVTGHEGNSRKNSSNGNPNYDIDHKKTISSAI-IN-HAFLONTVMKN 261  
Qy 456 CNYK-SNNEKSWNCTGFTTNKFPCTCEPPRQTLCLGRTYLLHRGHEEDY-----KEH 507  
Db 262 CNYKRRERDWDG-----NTKDDVCTIPDRYQLCMKELTNLVNNTDTNFRDITFRKLY 316  
Qy 508 LLGASIYEA-----QLLYKYYKEKENALCSIIQNSYADLADIIGSDIHKDYGYKK 559  
Db 317 LKREKLIYDAAVEGDLILKANNRY-NKD-----FCQDIRWSLGDGFDIIMGTDMEGISYKV 372  
Qy 560 MEENLNKV-NKDKCKNEESLKIPEKWDENKENVKMSAVLNK-----ETCKDYDK 612  
Db 373 VENNLRSIFGTDEKAQOR-----RKQWNESKQAIWTAMYSVVRKRLKGNPIWTKLNVA 427  
Qy 613 FQKIPQFLRPFKEWGDGDFCEKREKIYSPESFKVEC-----KKKQDCENTCKNKGSEY 665  
Db 428 VNEPQIYRWIREGRDYVSELPTVEV---QKLKEKCDGKINYTDKVKYKPPCQACKSY 484  
Qy 666 KKWIDLKSEYKQVDKYTKDKNNKQMYDNTDEVKNKEANV-----YLKESKECKQVNF 720  
Db 485 DQWTRKKNQWDLNPKFISVKNAB-----KVQTAGIVTPYDILKQELDEFNEVAFE 536  
Qy 721 DKIFNESYEVEDWCKCKDEKYLNEIKYPTKTHDIYDIDTFSTFG-DGTPI----- 772  
Db 537 NEI-NKRDGAYIELC-----VCSVEEAK-KNTQEVVTVNVDNAAKSQATNSNPIQPDVSS 589  
Qy 773 -----SINANINEQSGKDTNTGNSETSD-----SPVSHE-----PESDAAINV 812

Db 590 KAEPVPGDSTHGVN---SQDSTTGKAVTGDQNGQNPASDVORSIAESVSAKV 646  
Qy 813 E-----KLSGD-----ESSETRGILDINDPSTVNNVNEVHDASNTQGSV-----852  
Db 647 DPQKSVSKRSDDTASVTGIAEAGKENGAGNSRPSBSTVEANGSPGDDTVNSASIPVVSGE 706  
Qy 853 -----SNTSDITGHGESSLNRTNAQDIKIRSG-----NEQSDNQESSHSS 896  
Db 707 NPLVTPYNGLRHSKNSDSDGPAESMANPDSNK---GETGKGQNDMAKATKDSNSS 762  
Qy 897 DNGSGSLT-----ICQVPSEDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEG 944  
Db 763 DGTSSATGDTTDAVDREINKGVPEDRDKTVGSKDGGEDSANKDAATVVGEDRIRENSA 822  
Qy 945 FDSRD-----SENG-----RGDTTSN---THDVRRTNIV 971  
Db 823 GGSTNDRSKNDTEKNGASTPDSQSDATALSKTESLESTESGDRITNTDTSLEKNKG 882  
Qy 972 SERVNSHDFIRGMANNAHQVIT-----QIENGLIRGOEESAGNSVNYKDNPK 1023  
Db 883 KEKDLQKHPKSNDDTNEEPNSQDTTDAEGHDRDSIKND---KAERRKHMKKDTFTKNTN 939  
Qy 1024 RSNPSSN---DHKNIOEYNSRDKRVREIILKSKONKCNNEYSMEYCTYSDERNSSP 1080  
Db 940 SHLNSNNLSNGKLDIKEKYRDKATREDIILMSSVRKNNISLEYCNVEDKISS- 998  
Qy 1081 GPCSRERKXLCQISDYCLKYFNYSIYNCIKSEIKSPYKCFKSEG-QSSSIPYFAA 1139  
Db 999 NTCRSREKXLCSSISDFCLNYPVYSVEYLSCKMKEFEDPSYKCFPKGGFKDKTYFAA 1058  
Qy 1140 GGLVVLVILLSSASRKGNEEYDICESNIEATFEE-NNYLKLRSI-----FNQEVQE 1193  
Db 1059 AGAL--LILLIASRKMKNDSE-----EATNFEFEYCDNIHRIPLMPNNIEHQ 1108  
Qy 1194 TNISDYS 1200  
Db 1109 STPLDYS 1115

## RESULT 13

US-10-153-273-2

; Sequence 2, Application US/10153273

; Patent No. 6962987

## ; GENERAL INFORMATION:

; APPLICANT: Sim, Kim L.  
; Chitnis, Chetan  
; Miller, Louis H.  
; Peterson, David S.  
; Su, Xin-zhaun  
; Welles, Thomas E.

; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe Martens Olson &amp; Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: California

; COUNTRY: US

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/153,273

; FILING DATE: 21-May-2002

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/210,288

; FILING DATE: &lt;Unknown&gt;

; ATTORNEY/AGENT INFORMATION:  
; NAME: Fuller, Michael  
; REGISTRATION NUMBER: 36,516  
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1115 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Plasmodium vivax  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-153-273-2

Query Match 11.6%; Score 754; DB 2; Length 1115;  
Best Local Similarity 24.5%; Pred. No. 2.5e-43;  
Matches 296; Conservative 177; Mismatches 416; Indels 318; Gaps 57;

Qy 199 FKEIFLISVNTAKLLYNKEGKDPISFCNELNSFSDPFSSFIGDDMPGGTDRVKGY 258  
Db 22 YKDDFSIT-----LNYHEGKKYLIILKR-----KLEKANNRDVCNFF 59  
Qy 259 IN--TKPSDYVKEQVSKLNNIKKEWWE-KNKANLWNHMIWNHKGNIKSCAIIIPAEEPQI 316  
Db 60 LHPSQVNNVLLERTITLLECKNEYVKGNGYKL-----AKGH---HC-----VEEDNL 105  
Qy 317 NLWIKENWENFLMEKRLFLNTKDCVENKRYEACFGGCRLPSCS-SYTSFPMKSKTKQMEV 375  
Db 106 ERWLQGTNER-----RSEENIKYK--YGVTCLKIKYAQMNGKRSRILKES 149  
Qy 376 L-----TNLYKKNSGVDKNFLNDLFFKNNKNDLDDFFKNEKEYEDLDCRY-----423  
Db 150 IYGAHNFNGSNYMEGKGGDKTG-----BEKGEHKTDSKTDNGKANNLMLDYETSSN 204  
Qy 424 -----TATIIGKPLNGPAKN-----DVD-----IASQINVNDLRGF-----G 455  
Db 205 GQPAGTLDNLVLFVTVGHEGNSRKNSSNGGNPYDIDHKTTISSAI-IN--HAFQNTVMKN 261  
Qy 456 CNYK-SNNEKSNWCTGTFTNKPFGTCPPRRQTLCLGRTYLLHRGHEEDY-----KEH 507  
Db 262 CNYKRKRERDWDG-----NTKKDVCIPDRRYOLCKMELTNLVNNTDTNFRHDIRFKLY 316  
Qy 508 LLGASIVEA-----QLLKYYKEKENALCSIIQNSYADLADIIGSDIIDKYGKK 559  
Db 317 LKRKLTYDAAVEGDLKLNNRY-NKD---FCKDIRSLGDFGDIIMGDMEGIGYSKV 372  
Qy 560 MEENLKNV-NKDKRNEESLKIIFREKWWDENKENVKWSAVLKNK-----ETCKDYDK 612  
Db 373 VENNLRSIFGTDEKAQR-----RKQWNEKSAQIWTAMMYSVKYKRLKGNFICKLNV 427  
Qy 613 FQKIPQFLRWFKWGGDDFCEKREKIKYISFESFKVEK-----KKKDCDNTCNKCKSEY 665  
Db 428 VNIEPQIYRWIREWGRDYVSELPTFV---QKLEKCDGKINYTDKKYCKVPPCQNAKCY 484  
Qy 666 KKWIDLKXEYEQVDKYTKDNKNKMYDNIDEVKNEANV-----YLKESKECKOVNFD 720  
Db 485 DQWITRKKNQWDVLSNKFISVKNAB-----KVQTAGIVTVPYDILKQELDSFNEVAPE 536  
Qy 721 DKIFNSPNEYEDMCKCDBIKYLNBIKYPKTKHDIYDIDTFSDTFG-DGTPI-----772  
Db 537 NEI-NKRDGAYIELC-----VCSVEEAK-KYQTEVTVNVDNAKSAQTSNPNISQPVDS 589  
Qy 773 -----SINANINEQQSGKTSNTGNSSETSD-----SPVSHE-----PESDAAINV 812  
Db 590 KAEKVPDSTHGVN---SQDSTTGKAVTGDQNGQNPASDVORSIAESVSAKV 646  
Qy 813 E-----KLSGD-----ESSETRGILDINDPSTVNNVNEVHDASNTQGSV-----852



Best Local Similarity 20.8%; Pred. No. 4.5e-32;  
Matches 303; Conservative 192; Mismatches 466; Indels 499; Gaps 68;

Qy	56	NIYDSYEDYVVK-----LINSFVENKSVK-----KRSLSFI-----NNKTSYD- 96
Db	1118	NIISDKYKEIHEQAQMSVNSGTEASATKXHIHIDRNVIEFLSELYQQGKSNKSGTSD 1177
Qy	97	--IIPSYSVYRNDKFNLSNEINSGN---TSSNNFANTSEISIGKDNKOYTFIOK-RT 149
Db	1178	SAVIGTNTVTYEN-----VGAYLHDTGFDCCQSQNEFCD---EKSDGKDNKBYAFRDKPQD 1230
Qy	150	HLFACGKGR-----KSIKWICRENS----- 169
Db	1231	HDGACGCKSGSKPRVQIKTKKAAEKDTECKTVNDILKENDGKQVEDCHPKKNSGYP 1290
Qy	170	-----EKTYVCVPDRKIQLCVANFLN-----SRLETWEKFEKFLFLSVNTEA--KL 213
Db	1291	DWOCGNINLVEDPRVCMPPRQRQLCVHFLANDNEIKKLOQVNLKEAFIKSAAEATPFSW 1350
Qy	214	LYNKN-----EGKDPISPCVELNRNSFDSFRSFIGDDMDFG-GNTDRVKGYI 259
Db	1351	YYYSKDCGEGNLDKELKEKIPPAFLRSWFYTFGDRDLPFGTDISKHGEGSKLKEQI 1410
Qy	260	NTRFSDYYKEKNVEKLN---NIKEMWEKNKANLWNHMIYNHKGNTSKECAIIP----- 310
Db	1411	DSLFI-----KNGDQKSPNGKTRQEWWEHSHETWEAML-----CALVKIGAKKD 1454
Qy	311	-----ABEQIINLWIKENWENFLMEKGRULFNLIKOKCVENKK 347
Db	1455	DFTENGYNNVNFSDSKSTLTLEEPAKRQPLRWLTWYDDCYTRQYKLVQVBCKSNDQ 1514
Qy	348	YEAFGCGRLPCSSYTSFMKSKTQMEVLNLYKKNSGVDKKNFLNDLFKKNKNKDLDD 407
Db	1515	LK-CDTECNKCEDYVKMKKEWIP-QDKYKDER---DKGRF----- 1554
Qy	408	FFRKEEYDLDLCRYTATIIKSPFN-----GPAKNVDVIAQINVNDL----- 451
Db	1555	-----DRQHIGVWVTDYTGTTNATDYLNRKTASTCGDKPGSASV-QRNIQLLEQAYYDAD 1609
Qy	452	RGRGC-NYKSNNEKSNV-----CTG-----TPTNKKFPGTCE-----PP 483
Db	1610	KHCGCTKFIENDKYTNISSKDKCKGLVKEANTGAIKWQNKGNPNVNNLKELTEDVLFPS 1669
Qy	484	RROTLCIGRYTLHRRGH-----EEDYKEHLGASIEYAQLKYKYEKENAL----- 531
Db	1670	RLRIICF-----HALDGNITDPEVDKENGRLRMEVAATEGVNLOYKYEKKEKEKITS 1726
Qy	532	-----CSIIQNSYADLADIIKGSDIIDKYGKMEENLNKV-NKD-----KKRNEE 576
Db	1727	AHKYSVEVPFCSAMKYSFYDLRDLILGIDNLED-EKQKTEENLKKIFNKGNTSVGKSGDS 1785
Qy	577	SL-----KIFREKWDENKENVKNVMSAVLK-----NKETCKDYDFKQIP----- 617
Db	1786	TTGNPGSTARKFPFNNENKECVNNAMICGYKRRDRDGNSGNSARSDEDLKKGSPVSDDDY 1845
Qy	618	-----OPLRFWEKWDGDFCEKREKIYSFESFKVECKKCCDEN-----TCKN 660
Db	1846	PMGNRDEGTAYQPLRFWFAEWGDFC-KHKEK--ELEKLVGACNDYTCGDNEDKRXKCTD 1902
Qy	661	KCSEYKMWIDKKSSEYKQVDKYTKOKNKKYDNIDEVKNKE-ANVYLKESKECKDWNF 719
Db	1903	ACTQYKKFISWKPKQVEKQIKKYGENKD-KIYSEHPVKAQDAEDAREYLDQLK----- 1954
Qy	720	DDKIFNESPNEDCMKCCDEIKYLNIEIKPKTKHDIYDITDTSDETDGDTPTISNANIN 779
Db	1955	--KICENKSGDCBYCKWK-----DVSTQRLTDGNSQNNMPASLD 1990
Qy	780	EQSGGKDTNNTGSETSDSP--VSHEPESDAAINVEKLSGDESSSETRGILDINDP-SVT 836
Db	1991	DEP--KEVEGKCNQVPRGRPRVRETFPSRVSLSKATA--SKKEAKTAPPTKQPKKVE 2046
Qy	837	NNVNEVHDASNTQGSVSNTSDITNGHSESSLNRTTNAQDIKIGRSGNEQSDNQENSHS 896

```

;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-08-487-826B-14

```

Query Match 7.7%; Score 498.5; DB 1; Length 3060;  
Best Local Similarity 18.9%; Pred. No. 4.8e-25;  
Matches 317; Conservative 180; Mismatches 520; Indels 661;

Qy	90	NKTKSYDIIIPSPYSYRNDKFNLSJSENEDNGTNTNNNFANTSE--ISIGKDNKQYTFIQKR	148
Db	1309	NKXYTFQPPPEVATACDCNRSQTEPPKKKEENVBESACKIVEKILEGGK--	1360
Qy	149	THLPACGIGKRSIKWICRENS--KITVCVPDRKIQLCVANFLN-----SRLETWEKFKEI	202
Db	1361	TTVGECPKESYPDWDCNNNIDISHGACMPRPKQKCLYIIAHESQTEWIKITDDNLUKDA	1420
Qy	203	FLISV-----NTEKLLYNNKNEGDPISFICNELRNSPDSFRSSPIGDDMD	247
Db	1421	FIKTAATAETLSWQYYKSKNDSEAKIL--DRGLJPQFLRSMYTFGDIR-----D	1469
Qy	248	PGGNTDRVKGYINTKPSDYKBNVEKLNII-----KCBWWEKKNKANLNNHM	294
Db	1470	ICLNTD-----ISKQNDVAKAK--DKIGKFFSKDGSKPSGLSRQEWKWTNGPEIWKGM	1522
Qy	295	IV-----NHKGNISKECAIIPABEPOINLWIKENSNEL	328
Db	1523	LCALTKVYTDNDKTKINDYSDKVNQSONGNPSLEEF--AAKPOFLRMWIEWGEFC	1579
Qy	329	MEKKRLFLNIKDKVENKCYEACFGG--CRLPCSSYTSFMFKSKTQMEVLTN-----	378
Db	1580	AERQKKEENIKDACNEINSTQCNDAKHRCNQACRAOBYEVENKCKEFGOTNNFVLKAN	1639
Qy	379	-----LYKKNSGVDKNFLNDLFKGNKNNDLD-----DPFKNEKEYD	416
Db	1640	VQOPDPBYKGYEKGQVQPIQGNREYLLQKCDNNKSCMDGNVLSVSKPEKPFPGYIAHKYP	1699
Qy	417	DLDCRYTATIIKSFNLGPAKND-----VDIASQINV--NDLRGF--GCNYSKN	461
Db	1700	EKDCU--YQKHVPSIPPPPPVQOPAPVTVDVDCSIVKTLFKDTNNFSDACGLKYGKT	1758
Qy	462	NKSWNCTGTFTNKPGCT-----CEPPRQOTLCIGR-----	492
Db	1759	APSSWKIPSDTKSGAGATTGSGSGSGSICIPPPRRRLVYGLQEWATALPQEGGAAPS	1818
Qy	493	-----TYLLHRGHEEDYEKHLGASIIYEAQLLKYYKEKDENALCSI	534
Db	1819	HSRADDLRNAPIOSAAIJETTFFLDRYKEKKPQDGSQOALSQITS--TYSDDDEPDPKL	1877
Qy	535	IQN-----SYADLADIIGSDIIKDYGKKMBEENLNKVNKKRNEESLUKIF	581
Db	1878	LQNGKIPDPFLMLFYTLGDYRDLVHGGNTSD--SGNTYNSNNNNIVLEASGNEDHQKI	1936
Qy	582	RE-----KWDENKENVWKMVASVL-----KNEKFTC-----	607
Db	1937	QEKIEQILPKNGGTPLVPKSSAQTDPKWWEHABSIWKGMICALITYTEKPNPTSGARDEN	1996
Qy	608	---KD---YDKF-----QKI-----	616
Db	1997	KIEKODVEYKFFGSTADKHGTASTPTGYTKQVDYEVKLEDTSGAKTPSASDTPLLS	2056
Qy	617	-----POFLRWFKEWGDDFCBKRBKYSIPESFKVECKCK-----	651
Db	2057	DFVLPPYPFYLEBWGNFCKGRKHL--AQIKHECKVEBWGGSRGGITRQYSGDGE	2113
Qy	652	-----DCDENTCNKCSYKWKWDLKSEYKQVDKYTKOKNKKWYDNIDE	697
Db	2114	ACNEMLPKNDGTVPDLBKPSCAKPCSSYRKWIESKGKEFKQEKAYEQDKCKVNGS---	2170
Qy	698	VQKNEANVYLK-----EKSKE-----CKQVNF-----DDKIKFNESPN-----EYED	733
Db	2171	--NKHDNGFCETLTTSSKAKDFLTKLFGCKPNNVGEKTIIPDDDKTFKHTKDCDPCLPKFS	2228
Qy	734	MCKK--CDEIKYLNIEIKYPTKTHDIYDI-----DTPGSDTFGDGTPISINA	776

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israeleen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2710 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-08-568-459A-12

Query Match          7.6%; Score 495.5; DB 1; Length 2710;
Best Local Similarity 19.5%; Pred. No. 6.5e-25;
Matches 268; Conservative 170; Mismatches 482; Indels 453; Gaps 57;

QY 112 LSENEGSGTNNFANTSEISIGKDNKQVTFIQKTHLFCAGIK-RKSIKMICRENSE 170
DB 912 INENYSDSRNSGGPCT-----GKGDH-----GGVRIGTEWSNIEGKK 953
QY 171 KIT---VCVPRKIQLCVANFLNSRLFTM---EKFEIFLISVNTAKL----- 213
DB 954 QTSYKVNFLPRRHHMCTSNLENLDVGSVTKNDKASHLLGDVQLAAKTDAAEIKRYKD 1013
QY 214 -----LYNKEGKDPISFCNELRNSFSDPRSSFIGDDM---DFGNTDRVGYI--- 259
DB 1014 QNNIQLTDP1QQKQDEAMCAVRYSFADLGDITNGRDMWDEKSSDTMETRLITVFNKIK 1073
QY 260 -----NTKFSDYKKEKVKVEKLNNIKKEWKEKQKLANLWNHMINVNHKGNISKECAIIPA 311
DB 1074 EKHDGIKDNPKYTGDESKPAYK---KLRAWNEANRHOVWRMAMKCATKGII---CPGMPV 1128
QY 312 EE--PQINLMIKENENFLMEKKRLFLNIKDCKVE--NKKYEAC-----FGCRLPCSS 361
DB 1129 DDYIPQRLRWTEWAENYCAQSQEYDKLKKICADCMKGGDKCTQGDVDCGKCAACDK 1188
QY 362 YTSFMKSKTQMEVLTNLY-----KKKNSGVDKN-----NFLNDLPK--- 398
DB 1189 YKEIEKWNQWRKISKDKNLLYLQAKTTSTNPGRTVLGDDDDPDYQOMVDVFLTPHKASI 1248
QY 399 ----- 398
DB 1249 AARVLVKRAAGSPTEIAAAPTSTYTAAGYIHOEIGYGGCQEQTFCEKKHGATSTSTT 1308
QY 399 KNNKNDLDDFPKNEKYDDLDCCRATATIKSFLNGPAKNDVDIASINVND--LRGFG- 455
DB 1309 KENKE--YTFKQPPPEYATADC-----INRSQTEPEKKKEENVESACKIVEKILEGKNG 1361
QY 456 -----CNYKSNKSNKNCCTFTNKPFGTCPEPRRQTLCLGRYLLHRRHEEDYK--EH 507
DB 1362 RTTVGECNPK--ESYPDWDCNNIDISHDGAAMPFRQKLCI--YYIAHESQTEINIKTDDN 1418
QY 508 LLGASIVEA--QLLKVKYKEKBNALCSIIQN-----SYADLADLIKSGSDI 551
DB 1419 LKQAFIKTAAEFTLSQYYSKNDSEAKILDRGLIPSQFLRSMWYTFGDRICLNTDI 1478
QY 552 IKDYIG-KQWENLNK-VNKDKKNEESLKIIFREKKWDEKENENVKWMSAVL-----X 602
DB 1479 SKQNDVAKAKIGKGFPSKDGSKSPSGLS--RQEWKNTNGPEIWKGMCLCALTXYVTDTD 1536
QY 603 NKETCKD---YDK-----FQKIPQLRWFKNGDDDFCEKRKEKIYSPFSFKV 646
```

```
DB 1537 NKRKIKNDYSYDKVNQSONGNPSLEBFAAKPQFLRWMIENGEEFCAERQKK---ENIIKD 1593
QY 647 ECKK----KDCDB--NTCKNKSEYKKWIDLKSEYEKQVDKYT-----KDKNKQMYD 693
DB 1594 ACNEINSTQCCNDAKHRCNQACEAYQYEVYENKKKEFGSGQTNNFVLKANVQPDPEYKGYE 1653
QY 694 NIDEVKNKEANVYLKEKSECKDVNFDKIFNESPNE-----YEDMKCKCDEIKYL 744
DB 1654 YKGVQPIQGNBYLLQCDNNKSCMDGNVLSVSPKEKPGKYAHKYPEKC-DCYQKGHV 1712
QY 745 NEIKYPK-----TKHDIYDIDTFSDT-----FGDGTPIISINANINEQ 781
DB 1713 PSIPPPPPVQPOPEAPTVTVDVCSIVKTLFKDTNNFSDACGLKYGKTAPSSWKCPISDT 1772
QY 782 QSGKOTSNKTSNSETSDSPV-----SHEPESD----- 807
DB 1773 KSGAG-ATTGKSGSDSGSICIPRRRLRYVGLQEWATALPQEGGAAPSHSRADDLNAP 1831
QY 808 ---AAINV-----EKLGDSESSSETRGIL-----DINDPSVTNNVNEVHDASNTQ 849
DB 1832 IQSAAIETFTFLWDRYKEKKPGQDGSQALSQLTSTVSDDEEDPP-----DKLLQN 1882
QY 850 GSV-----SNTSDI---TNGHSESSLNRTTNAQDIKIGRSNEQ 885
DB 1883 GKIPDPFLRLMFYTLGDRDILVHGNTSDSGNTNG-----SNNNNIVLEASGNKE 1933
QY 886 SDN--QENSSHSDNSGSLTI-----GOVPSE---DNTON-----TYDSQNERDT 926
DB 1934 DMQIOEKEIQLPKNGGTPLPKSSAQTDPKWNNEHAEISIKGMCALTYTEKNPD TSA 1993
QY 927 PNALASLPDDDKINEIEGFDSSRDSSENGRDTTSTNTHDVRRTNIVSERKVNSHDFIRNGM 986
DB 1994 RGDENKIEKXDEV--YEKFFGSGTADKHGTASTPTGT----- 2027
QY 987 ANNAHQAQITQIENNGIIRGOEESAGNSVYKNDPKRNFSSSENHKKNIQEYNSRDTK 1046
DB 2028 -----YKTOYDYEKVKLEDTSGAKTPSASDTPFLSLDFVLRPPFYRYLEWGMQNFCK 2079
QY 1047 RVREEIILKSKONK-----CNNEYSMEYCTYSD-ERNSSPGP 1082
DB 2080 KRKGKLAQIKHECKVEENGSGGSRGGITRQYSGDGEACHEMLPKNDGTVDPDLKPCSAKP 2139
QY 1083 CSR-----BERKKLCCQISDYCLKYENFYNSIYNYCNICKSEIKSPY 1123
DB 2140 CSSYRKWIBSKGFEKQEKAYEQKDKCVNGSNKHDNGFCETLTSSKAKDF 2192
```

## RESULT 18

```
US-08-487-826B-12
; Sequence 12, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B

FILING DATE: 10-SEP-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CP1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 2710 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium falciparum

US-08-487-826B-12

Query Match 7.6%; Score 495.5; DB 1; Length 2710;

Best Local Similarity 19.5%; Pred. No. 6.5e-25;

Matches 268; Conservative 170; Mismatches 482; Indels 453; Gaps 57;

```
QY 112 LSENEISGNTNSNNFANTSISIGKDNKQYTFQKRTHLFAGIK-RKSIKWTCRENS 170
DB 912 INENYSNDRSGNGPCT-----GKGDH-----GGVRMRIGTEWSNIEGKK 953
QY 171 KIT---VCVPRKIQLCVANFLNRLTMM---EKFKIFILISVNTAKL----- 213
DB 954 QTSYKNVFLPRRHEMCTSNLENLDVGSVTKNDKASHLLGDVLAQKTDAAEIIKRYKD 1013
QY 214 -----LYNKNKGKDPISFCNBLRNSFDSFSSFIGDDM--DFGNGTNRVKGYI----- 259
DB 1014 QNNIQLTDPIQKQKQAMCRAVRYSFADLDGIIRGRDWMDEKSSDTMETRLITVFKNIK 1073
QY 260 -----NTKFSYKKEKVEKNLNNIKKEWEKQKANKLNNHMIVNHKGNISKECALIPA 311
DB 1074 EKHDGIDKNPKYTGDSEKPKPAYK--KLADWWEANRHQVWRAMCATKGII---CPGMVPV 1128
QY 312 EE--POINLWIKWENFLMEKKRLFLNLIKDKVE--NKKEAC-----FGCRLPCSS 361
DB 1129 DDYIPQRLRWTEWAEWYCKAQSQR-----DKLKKICADCKSGDKGCTQGDDVDCGCKAAADK 1188
QY 362 YTSFMKSKTQMEVLTNLY-----KKNNSGVDKN-----NPLNDLFK--- 398
DB 1189 YKEIEKWNQWRKISDKYNLLYLAQKTTSTNPGRTVLGDDDDPDYQQWVDPLTPIHKASI 1248
QY 399 ----- 398
DB 1249 AARVLVKRAAGSPTEIAAAAPITPYSTAAGYIHOEIGYGCQEQTFQCEKKGATSTSTT 1308
QY 399 KNNKNDLDDPFKEKEDDLCDCRYTATIIKSLFLNGPAKNDVDIASQINVD--LRGFG- 455
DB 1309 KENK--YTFKQPPPEYATACD-----INRSQTEPKKKEENVESACKIVEKILEGKNG 1361
QY 456 -----CNYSNNKSNKNCCTFTNPKFGTCPPRROTLCIGRTYLLHRGHEEDYK--EH 507
DB 1362 RTTVGECNPK--ESYPDWDCKNNIDISHGACWMPRRQKLC--YYIAHESQTENIKTDN 1418
QY 508 LLGASIVEA---QLLKVKYKKEBNALCSIIQN-----SYADIADIIGSDI 551
DB 1419 LKDAFIKTAATAETFLSWOYKSKNDSEAKILDRGLIPSQFLRSMWYTFGDRDCLNTDI 1478
QY 552 IKDYKG--KMEENLNK-VNKDKKNEESLKIIFREKWNENKENVKWSAVL-----K 602
DB 1479 SKQNDVAKADKTKGKFFSKDGSKSPGLS--ROEWKTKNGPEIWKGLMCLTKYVTDTD 1536
QY 603 NKETCKD---YDK-----FQKIPQFLRWFKWGDGDFCEKREKIKYFSFVKV 646
```

```
DB 1537 NKRIKNDYSDYKVNQSGNGNPSLEEAFAKQFLRWMEWGEERFAERQKK---ENIKD 1593
QY 647 ECKK-----KDCDE--NTCKKCKSEYKWKWIDLKSEYEKQVDKYT-----KDKNKQYD 693
DB 1594 ACNEINSTQOCNDAKHRCNOACRAYQYEVENKKKEFGQTNNFVLKANVQPDPEYKGYE 1653
QY 694 NIDEVQKNEANVYLKSEKCKOVNFDKIFNESPNE-----YEDMCKCKDEIKYL 744
DB 1654 YKDGVPQIQGNEYLLQKDNCKSCMDGNVLSVSPKEKPFQKYAHKYPEK--DCYQGRHV 1712
QY 745 NEIKYPK-----TYKHDYIDITDTSDT-----FGDGTPIISINANINEQ 781
DB 1713 PSIPPPPPVQPOPEAPTIVTVVDCSIVKTLTKDTNNFSDACGLKYGTAPSWKCIPSDT 1772
QY 782 QSGKDTNTGNSSETSDSPV-----SHEPESD----- 807
DB 1773 KSGAG-ATTGKSGSDSGSICIPPRRRRLYVGLQEWATALPQGEAAPSRSRADDLRNAP 1831
QY 808 ---AAINV-----EKLSGDESSESTRGIL-----DINDPSVTNNVNVHDASNTQ 849
DB 1832 IQSAAIETFFLWDRYKEEKKEKPGDGSQOALSQLTSTYSDDDEDP-----DKLLQN 1882
QY 850 GSV-----SNTSDI--TNHSESSLNRTTNAQDIKIGRSGNEQ 885
DB 1883 GKIPDPFLRMFYTLGDYRDILVHGGNTSDSGNTNG-----SNNNNIVLEASGNKE 1933
QY 886 SDN--QENSHSSDNSLSLTI-----QGVPE-----DNTQN-----TYDSQNPHRDT 926
DB 1934 DMQIKQIKBIQILPKNGGTFLVPKSSAOTPDKNWNEHAESIWKGMICALTYTEKNPDTSA 1993
QY 927 PNALASIPSDDKINEIEGFDSRSSENGRGDTTNTHDVRRTNIVSERRVNSHDFIRNGM 986
DB 1994 RGDEKIEKDEV--YEKFGSTADKHGTASTPTGT----- 2027
QY 987 ANNAHHQYITQIENNGIIRGQEBESAGNSVNYKONPKRSPSSNDHKHKNIOEYNSRDTK 1046
DB 2028 -----YKTQVDYKVKLEDTSAGKTPSASSDTPLLSDFVLPRFPYFRYLEEWGQNECK 2079
QY 1047 RVREEIILKSKQNK-----CNNEYSMEYCTYSD--ERNSSPCP 1082
DB 2080 KRHKLAQIHKCEKVEENGSGRRGGITRQYSGDGEACNEMLPKNDGTVPDLKPSCAKP 2139
QY 1083 CSR-----BERKKLCCQISDYCLKYFNFYISYVNCIKSEIKSPY 1123
DB 2140 CSSYRWKIESKGKEFEKQKAYEQKCKVNGSKNDHNGFCETILTSSKADF 2192
```

# RESULT 19

US-09-210-288-12

Sequence 12, Application US/09210288

Patent No. 6392026

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellem, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/210,288
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH21.1FPDVI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2710 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORGANISM: Plasmodium falciparum
; US-09-210-288-12

Query Match 7.68; Score 495.5; DB 2; Length 2710;
Best Local Similarity 19.54; Pred. No. 6.5e-25;
Matches 268; Conservative 170; Mismatches 482; Indels 453; Gaps 57;

Qy 112 LSENEGSGTNNFNANTSEISGKDNQKQVTFQKTHLPACGK-RKSKWICRENS 170
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 171 KIT---VCVPRKIQLCVAFNLSRLTM---EKKEIFLISVNTAKL----- 213
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 954 QTSYKVFPLPRRHHMCTNLENLDVGSVTKNDKASHLLGDVQLAAKTADEAEIKRYKD 1013
Qy 214 -----LYNKEGKDPISFCNELRNSFSDFRSSFTGDDM--DFGNTDRVKGVI----- 259
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1014 QNNIQLTDPIQCKQDQAMCAVRYSFADLGIIRGDMWDEKSSSTDMETRLITVFRNIK 1073
Qy 260 -----NTKFSYVYKKNVEKLNKIKCEWWEKKNKANLWNHMTVNHKGNISKECALIPA 311
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1074 EKHGIGKDNPKYTGDESKKPAYK--KLRAADWEANRHOVWAMKCATKGI--CPGMPV 1128
Qy 312 EE--PQINLMIKEWENFLMEKRLFLNIDKQVVE--NKYEAC-----FGGRLPCSS 361
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1129 DDYIPQRLRWTEWAEWYCAQSQEYDKLKKI CADQMSKGDGKCTQGDVDCGKCAACDK 1188
Qy 362 YTSFPMKSKTQMEVLTNLY-----KKKNSGVDKN-----NFLNDLFK--- 398
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1189 YKEBIEKWEQWRKISDKYNLLYLQAKTTSTNPGRTVLGDDDDPDYQQMVDPLTPHKA 1248
Qy 399 -----AARVLVKRAAGSPTEIAAARIPYSTAAGYIHOEIGYGGQEQTPQCEKHGATST 1308
Qy 1249 AARVLVKRAAGSPTEIAAARIPYSTAAGYIHOEIGYGGQEQTPQCEKHGATST 1308
Qy 399 KKNKNDLDDFPKBEKYDDLDCCRYTATIKSFLNGPAKNDVDIASQINVD--LRGFG- 455
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1309 KENKE--YTFKQPPPEYATADC-----INRSQTEPKKKEENVESACKIVEKILEGKNG 1361
Qy 456 -----CNYKSNKSNKNTCTFTNKPFGTCPEPRRTCLGRYLLHRGHEEDYK--EH 507
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1362 RTTVGECPNK--ESYVPDCKNNDIISHDGACMPPRKOKLCI--YYIAHESQTEINIKTDDN 1418
Qy 508 LLGASIVIEA---QLLKVKYKEDENALCSIIQNT-----SYADLADIIGKSDI 551
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1419 LKDAFIKTAARETFLSWQYKSKNDSEAKILDRGLIPSOFLSRMMYTFGDRYDCLNTDI 1478
Qy 552 IKOYVG--KQMEENLNK--VNKDKKNESLSKITFRKKNWDEKENYVKNVSAVL-----K 602
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1479 SKQNDVAKADKIGKFPKDGKSPSGLS--RQEWKTKNGPEIWKGMALCALTKYVTDTD 1536
Qy 603 NKETCKD---YDK-----FQKIQFLRWFKWEGDDDFCEKREKIEYFSEPKV 646
```

```
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1537 NKKIKNDISYDKVKNOSQNGNPSLEBFAAKPQLRWMIEWGEFCAERQKK--ENIKD 1593
Qy 647 ECKK---KDCDE--NTCKNKSEYKKWIDLKKSEYEKQVDKYT-----KDKNKKMYD 693
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1594 ACNEINSTQOCNDKAKHRCQACRAYQYEVVENKKEFSGQTNNFVLKANVQPDPEYKGYE 1653
Qy 694 NIDEVKNKEANVYLKESKECKDVNFPDDKIFNESPNE-----YEDMCKCKDEIKYL 744
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1654 YKGVQPIQGNVEYLLQCKDNKNCSCMDGNVLSVSPKEKPFQYAHKYPERC--DCYQKGV 1712
Qy 745 NEIKYPK-----TKHDIYDIDTFSDT-----FGDGTPIISANINEQ 781
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1713 PSIPPPPPVQPOPEAPTVTVDVCSIVKTLFKDTNNPFSACGLKYGTATSSWKCPISDT 1772
Qy 782 QSGKDTNTSGNSTSDSPV-----SHEPESD----- 807
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1773 KSGAG--ATTGSGSDSGSICIPPRRRRLVYGLQEWATALPQGEGAAPSHSRADDLRNF 1831
Qy 808 ---AANV-----EKLSDGESSETRGIL-----DINDPSVTNNVNEVHDASNTQ 849
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1832 IQSAAIETFFLMDRYKEKKPQDGSQALSQLTSTYSDDDEDP-----DKLLQN 1882
Qy 850 GSV-----SNTSDI---TNGHSESSLNRTTNAQDIKIGRSGNEQ 885
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1883 GKIPDPFLRMFTLGDYRDILVHGNTSDSGNTG-----SNNNNIVLEASNGKE 1933
Qy 886 SDN--QENSSHSDNSGSLTI-----GQVPESE---DNTQN-----TYDSQNPHRDT 926
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1934 DMQKIOEKIEQLPKNGGTPLVKPSAQTPDKWNEHAESIKWGMICALITYTEKNPDTSA 1993
Qy 927 PNALSLPDDKINIEIGFDSSRDSSENGRGDTTNTNHDVVRTNIVSERRVNSHDFIRNGM 986
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 1994 RGDENKIEKDDVE--YEKFFGSTADKHGTAFTGT----- 2027
Qy 987 ANNAHQYITOLENNGIIRGOESAGNSVYKDNPKRNSFSENDDHKKNIQEVNSRDTK 1046
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 2028 -----YKTIQDYBEKVKLEDTSGAKTPSASDTPLLSDFLVRPFRYLEEWGQNFCK 2079
Qy 1047 RVREEIILKLSKQNK-----CNNEYSMEYCTYSD--ERNSSPGP 1082
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 2080 KRXHLAQIKHECKVEENGSGSRGGITRQYSGDGEACNEMLPKNDGTVPDLEKPSCAKP 2139
Qy 1083 CSR-----BERKKLCCQISDYCLKYNFYSIEYNCIKSEIKSPY 1123
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 2140 CSSYRKWIBSKGFEKQEKAYEQKDKCVNGSNKHDNGFCETLTITSSKADF 2192
```

## RESULT 20

US-10-153-273-12  
Sequence 12, Application US/10153273  
Patent No. 6962987

## GENERAL INFORMATION:

APPLICANT: Sim, Kim L.  
Chitnis, Chetan  
Miller, Louis H.  
Peterson, David S.  
Su, Xin-zhaun  
Welless, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe Martens Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: California  
COUNTRY: US  
ZIP: 92660

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/153,273

FILING DATE: 21-May-2002

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/210,288

FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Fuller, Michael

REGISTRATION NUMBER: 36,516

REFERENCE/DOCKET NUMBER: NIH121.1FWDV1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 2710 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORGANISM: Plasmodium falciparum

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-153-273-12

Query Match 7.6%; Score 495.5; DB 2; Length 2710;

Best Local Similarity 19.5%; Pred. No. 6.5e-25;

Matches 268; Conservative 170; Mismatches 482; Indels 453; Gaps 57;

Qy 112 LSENNDSNGTNNFANTSBISIGKQNKQYTFQKRTFLPACGK-RKSIKWLRENSE 170

Db 912 INEYNSDRSGSGPCT-----GKGDH-----GGVRMRIGTENSIEGKK 953

Qy 171 KIT---VCVPDRKIQLCVANFLNRLSTM---EKFEIFLISVNTAKL----- 213

Db 954 QTSKVNFLPPRRHEMCTSNLENDVGSVTXNDKASHSLGDLQVLAAKTDAAEIIRYKD 1013

Qy 214 -----LYNKGKDPISFCNLRNLSFDSFSFTGDDM--DPGNTDRVKYI----- 259

Db 1014 QNNIQLTDPIQKQOEAMCRAVRYSFADLGIIRGRDWDKSDTDMETRLITVFRNIK 1073

Qy 260 -----NTKFSDYKKEKVKELNNIKKEWKEKKNLNNHMLVNHKGNISKCAIIPA 311

Db 1074 EKHGIGKNDPKYTGDSEKPAYK--KLRRADWEANRHQVWRAMKATGII---CPGMPV 1128

Qy 312 EE--PQINLWIKENENFLMEKGRFLNLIKDKCQE--NKYEAC-----FGGRLPCSS 361

Db 1129 DDYIPQRLRWTEAWEYCKAQSQEYDKLKKICADCKSKGDKCTQGDDVDCGCKAACDK 1188

Qy 362 YTSFKSKSTQNEVLNLY-----KKNSGVDKN-----NFLNDFK--- 398

Db 1189 YKEBIEKWNEQWRKISDKNLLYLAQATTTSTNPGRTVLGDDDDPDYQQWVDFLTIHKASI 1248

Qy 399 ----- 398

Db 1249 AARVLVRAAGSPTEIAAAAPITPVSTAAGYIHOEIGYGGCQEQOTQCEKKGATSTTT 1308

Qy 399 KNNKNDLDDFFKNEKEYDDLCDCTRYATIIKSLFNGPAKNDVDIASQINVND--LRPG- 455

Db 1309 KENKE--YTFKQPPEYATADC-----INRSQTEEPKKEENVESACKIVEKLEGKNG 1361

Qy 456 -----QNYKNNKSNWCTGFTNKNPGTCEPRRQTLCLGRYLLHRGHEEDYK--EH 507

Db 1362 RTTVGECNPK-ESYPDMWDCKNNIDISHGACWPPRRQKCL--YYIAHESQTEINIKTDDN 1418

Qy 508 LLGASIIYEA---QLKYYKKEKENALCSLTON-----SVADLADIKGSDI 551

Db 1419 LKDAFIKTAABETFLSQYKNSDKSEAKILDRGLPSQFLRSMYTFGDIYRDLCLNTDI 1478

Qy 552 IKDYYG-KMBENLNK-VNKKRNEESLKFREKWDENKENYKVMASVL-----K 602

Db 1479 SKQNDVAKAKDKIGKFFSKDGSKSPGLS--RQEMWKTNGPEIWKGMLCALTKYVTD 1536

Qy 603 NKETCKD---YDK-----FQKIQFLRWFKWEGDDDCRCEKEKESYSESFKV 646

Db 1537 NKRKIKNDYSYDKVNSQNGNPSLEEFAAKQFLRWMIIEWEGEEFCAERQKK--ENIKD 1593

Qy 647 ECKK---KDCDE--NTCKNKCSEYKWKWIDLKSEYEKQVDKYT-----KDKKKWYD 693

Db 1594 ACNEINSTQOCNDAKHRCNQACRAYQEVVENKKEPFGQTNNFVLKANVQDPPEYKGYE 1653

Qy 694 NIDEVKKNKANVYLKESKECKQVNFDDKIFNBSGNE-----YEDMCKKCDDEIKYL 744

Db 1654 YKDGVPQIQGNEYLLQKCDNNKSCMDGNVLSVSPKEKPFQKXAHKYPEKC--DCYQGHV 1712

Qy 745 NEIKYBK-----TKHDIYDIDTSDT-----FGDGTPTSINANINEQ 781

Db 1713 PSIPPPPPVQVQPEAPFTVTVDVCSI VKTLFKDTNNFSDACGLKYGTAPSWKCPISDT 1772

Qy 782 QSGKDTNNTGNSSETSDSPV-----SHEPESD----- 807

Db 1773 KSGAG-ATTGKSGSDSGSICIPPRRRRLYVCKLQEWATALPQGEGAAPSHSRADDLNRAF 1831

Qy 808 ---AAINV-----EKLSDSESSETRGIL-----DINDPSVTNNVNEVHDASNTQ 849

Db 1832 IQSAAIETFFLWDYRKEBKPKQGDGSOALSQLTSTYSDDDEDP-----DKLLQN 1882

Qy 850 GSV-----SNTSDI--TNGHSESSLNRTTNAQDIKIGSGNEQ 885

Db 1883 GKIPDFLRLMFTGLDYRDILVHGNTSDSGNTNG-----SNNNNIVLEASGNKE 1933

Qy 886 SDN--QENSHSDNSGSLTI---GQVPE---DNTON-----TYDSQNPHRDT 926

Db 1934 DMQIKQIEQILPKNGGTLVPKSSAGTQDKWNEHAESIWKGMICALTTEKKNPDTSA 1993

Qy 927 PNALASLPDDKINEIEGFDSSRDSSENGRGDTTNTNHDVRRNTIVSERRVNSHDFIRNGM 986

Db 1994 RGDENKIEKDEV--YEKFFGSTADKHGTASTPTGT----- 2027

Qy 987 ANNAHAHQYITQIENNGIIRGOESAGNSVNYKDNPKRNSSESSNDHKKNQOYNSRDTK 1046

Db 2028 -----YKTQDYKVKLEDTSGAKTPSASDTPLLSDFVLPPPYFRYLEEWGQNFCK 2079

Qy 1047 RVREIILKSKONK-----CNNEYSMEYCTYSD--ERNSSSPG 1082

Db 2080 KRKHKLAIKHECKVEENGSGRRGGITRQYSGGGEACNEMLPKNDGTVPDLKPSCAKP 2139

Qy 1083 CSR-----BERKLCQISDYCLKYFNFIYIYCNKSEIKSPY 1123

Db 2140 CSSYRKWIESKGEFEKQKAYEQKDKCVNGSKNDHNGFCETLTSSKAKDF 2192

RESULT 21

US-08-929-329-5

; Sequence 5, Application US/08929329

; Patent No. 6120770

; GENERAL INFORMATION:

; APPLICANT: Adams, John H

; APPLICANT: Dalton, John P

; APPLICANT: Kappe, Stefan

; TITLE OF INVENTION: Plasmodium Proteins Useful for Preparing

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Barnes & Thornburg

; STREET: 11 S Meridian

; CITY: Indianapolis

; STATE: Indiana

; COUNTRY: USA

; ZIP: 46204

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,329
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Breen, John P
; REGISTRATION NUMBER: 38,833
; REFERENCE/DOCKET NUMBER: 835910-28685
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 231-7745
; TELEFAX: (317) 231-7433
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1507 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium yoelii
; US-08-929-329-5

Query Match 6.7%; Score 432; DB 2; Length 1507;
Best Local Similarity 19.6%; Pred. No. 6.9e-21;
Matches 263; Conservative 161; Mismatches 403; Indels 504; Gaps 55;

QY 33 NRODESSDISRV-----NSPELNHNKHTNIYDSYEDVNNKL-----INSPFVENSVKK 81
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 436 QNIEYPDIEVNVKIWNKRNSELSNLYNDKKFKNSDINKGMAINMNDINEIKENSKLQT 495

QY 82 KRSLSFNNKTK-----SYDIIPPSYSYANDK-----PNSISENEDNS-GWYNS-NNF 127
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 496 NK-----GNETKTKTKYGLYPIPTISYLOIHKKMELKYNMDSNSTSFHNTNAPTHY 550

QY 128 ANTSEISIGKDNKYT-----FLNEIRT-----FIOKRTHLFACGI 156
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 551 EGNKSTFGVNNKRENTYGTQDINLRNNYNNQPKNPQAEYMDRFDIEKHYY----- 606

QY 157 KRKSIAK-----ICRENSEKITVCVPDRKIQLCVANFLNSRLB-TMEKFK 200
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 607 ----IDWKQDGKYGSGKLYKNIISHETADTIQSLLITDKDDICPNHYPGQAQSCPNYG 662

QY 201 EIFLI-----SVNTEAKLLYNKQEKQDSIFCNELRNFSDFRSFIDGDDMDFGNTDRVK 256
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 663 KSIIVKTPESINGNEHL--NSN-----FLNEIRT----- 689

QY 257 GYINTKFSYVYKKNVE-----KLNNIKKEWMEKN--KANLWNH----- 293
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 690 GYLNK-----YMKSNVZLPYEKSGLAMHGDLSVCPKSWDEENLYKQRDYNDYDMCKSTV 744

QY 294 -----MIVNHK-----GNISKECAIIPAEBPQINLIWKENNFNM 329
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 745 MKSTIPLKMFYKTKLLYGLYGLGRLGNSNISKVKNIFKSPQNNITL-----PM 795

QY 330 EKKGLFLNKD-----KCVENKYEACFGGCRLPCCSYTSFMKKSXTQMEVL----- 376
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 796 FNPSSIKNLLDCSLYSYLCGLPCLENAYNKCF-----RSLPAYFNHETNECIIILGTHE 848

QY 377 ---TNLYKKNSGVDDKNFNDLPEKN---NKNLDLDPFKNEKEYDDLCDCRYTAII-- 427
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 849 QERNNNCKTRSDTDKNCQR--VRKNISTKNWTVYTSFIR--PDYBEKPPRYPLPKFS 904

QY 428 -----IKSPLNGPAKNDVDI----- 442
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 905 FGKYDEETGCKCSLIN--KKNIINILPSSCLEYMFIMYPSVLQRTKKNYGVWVASES 962

QY 443 ASQINVNDLRGFGCNYKNSNEKSNWCTGTFTNKFPGTCEPPRRPTCLGRTYLLHRGHEE 502
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 963 VNSSLNLYNAGK-ECY--INEPK-NCVIDKYNHF-----SFTSLTTNDI 1002
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```





```

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 921 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
;
US-09-210-288-8

Query Match          5.6%; Score 360.5; DB 2; Length 921;
Best Local Similarity 22.4%; Pred. No. 3e-16;
Matches 166; Conservative 83; Mismatches 231; Indels 261; Gaps 33;

QY 217 KN-EGKPSIFCNELRNFSDFRSFSGDMDPGGNTDRVKGYNITKFSDDYKKN-----271
DB 58 KNCNGGDTGKCG-SNGKNGNDYITASDIENGNS---IGNIDMVVSD--KDANGFNG 111
QY 272 -----VEKLNKIKKEWKK-ANLWNHVIWKNHGNISKECAIIPAEPO---INLWIK 321
DB 112 LDACGSANIPKIRKEQKCAKVGCLDVCGLKNNGSIDK-----DQKQIIIRALLK 164
QY 322 EWNENFLMEKKRLFLNIDKCVENKYEACFGGRLPCSSVTSFMKSKTQMEVLTNLYK 381
DB 165 RWVEYFLDYNNKINAKI-SHCTKDNSTCTNDCPNKCTCVEEWINQKRTWKNIKKHYK 223
QY 382 KKNSGVDKN--NFLNDLF-----KKNK-----NDLDDFFKNEKEYD 416
DB 224 TONENGDNMKSIVTDILGALQPOSVDVNAIKPCSGLTAFESFCGLNGADNSEKGEEDY 283
QY 417 DL-----CDRYTATIISFLNG-----434
DB 284 DLVLCMLKNLEKIQEQCKKKHGETSVE---NGGKSCPTLDNTTLEEPFEEBENQVEAPNI 340
QY 435 -----PA-----KNDVDITAS 444
DB 341 CPKQTVEDKKEEBEETCTPASPEKPVPHVWRWTFPPVFKIWRGRNKTCTEIVA 400
QY 445 QINVDNLG---FGNYSKNNEKSNCTGTTNKP-----GTCEPPRRQTLC-----490
DB 401 EM-LKDKNGRTTVGECYRKETYSEWTC---ESKIKMGQHGACIPPRRQKLCLEHYLEKIM 456
QY 491 -----GRYLLA-----HRGHEEDYKEHLGASIEAQLLYKYEKDE 528
DB 457 TTNELKVAPIKCAAAETFLWQNYKDKNGNAEDLDEKLGKGIIPEDFKOMFY-----511
QY 529 NALCSIIQNSYADLADIIGSDIINDYYGKMEENLANKV--NKDKRNEESLKI-FREKW 585
DB 512 -----TPADYEDICLGTDISK---KDTSGVGVKVCNIDDVFKISIRYKSW 559
QY 586 WDKENKENVKMSAVLK-----NKETCK-----DYKFKQKIP 617
DB 560 WETNGPVWEGMLCALSVDTSLNINNVNPTHKLTGEGNNNFEKVI FGSDSSTLSKFSERP 619
QY 618 QPLRFKEGWDGFCRKEKELYSPEKVECK-KDCDEN-----TKNKCSEYK 666
DB 620 QFLRLTWGSENFCKEKKK-----YKVLAKCKDCVDGDKNGKVCVACKQCKQYH 673
QY 667 KWLIDKKSEYKQVDKYTKDKKKMYNDNIDEVQNK-EANVYLKESKSECK-----DWNF 719
DB 674 SWIGIWDNYYKQGRYTEVKKIPLYKEDKDKVNSDDARDLYLTQLQNMKCVNGTTDENC 733
QY 720 DDKIFNESPNEVDCKKCKDE 740
DB 734 EYKCMHKTSTNSDMPESLDE 754

RESULT 26
US-10-153-273-8
; Sequence 8, Application US/10153273
; Patent No. 6962987

; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; Chitnis, Chetan
; Miller, Louis H.
; Peterson, David S.
; Su, Xin-zhaun
; Wellens, Thomas E.
;
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
;
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/153.273
; FILING DATE: 21-May-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/210.288
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36.516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDVI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 921 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
;
US-10-153-273-8

Query Match          5.6%; Score 360.5; DB 2; Length 921;
Best Local Similarity 22.4%; Pred. No. 3e-16;
Matches 166; Conservative 83; Mismatches 231; Indels 261; Gaps 33;

QY 217 KN-EGKPSIFCNELRNFSDFRSFSGDMDPGGNTDRVKGYNITKFSDDYKKN-----271
DB 58 KNCNGGDTGKCG-SNGKNGNDYITASDIENGNS---IGNIDMVVSD--KDANGFNG 111
QY 272 -----VEKLNKIKKEWKK-ANLWNHVIWKNHGNISKECAIIPAEPO---INLWIK 321
DB 112 LDACGSANIPKIRKEQKCAKVGCLDVCGLKNNGSIDK-----DQKQIIIRALLK 164
QY 322 EWNENFLMEKKRLFLNIDKCVENKYEACFGGRLPCSSVTSFMKSKTQMEVLTNLYK 381
DB 165 RWVEYFLDYNNKINAKI-SHCTKDNSTCTNDCPNKCTCVEEWINQKRTWKNIKKHYK 223
QY 382 KKNSGVDKN--NFLNDLF-----KKNK-----NDLDDFFKNEKEYD 416
DB 224 TONENGDNMKSIVTDILGALQPOSVDVNAIKPCSGLTAFESFCGLNGADNSEKGEEDY 283
QY 417 DL-----CDRYTATIISFLNG-----434
DB 284 DLVLCMLKNLEKIQEQCKKKHGETSVE---NGGKSCPTLDNTTLEEPFEEBENQVEAPNI 340
QY 435 -----PA-----KNDVDITAS 444
DB 341 CPKQTVEDKKEEBEETCTPASPEKPVPHVWRWTFPPVFKIWRGRNKTCTEIVA 400
QY 445 QINVDNLG---FGNYSKNNEKSNCTGTTNKP-----GTCEPPRRQTLC-----490
DB 401 EM-LKDKNGRTTVGECYRKETYSEWTC---ESKIKMGQHGACIPPRRQKLCLEHYLEKIM 456
QY 491 -----GRYLLA-----HRGHEEDYKEHLGASIEAQLLYKYEKDE 528
DB 457 TTNELKVAPIKCAAAETFLWQNYKDKNGNAEDLDEKLGKGIIPEDFKOMFY-----511
QY 529 NALCSIIQNSYADLADIIGSDIINDYYGKMEENLANKV--NKDKRNEESLKI-FREKW 585
DB 512 -----TPADYEDICLGTDISK---KDTSGVGVKVCNIDDVFKISIRYKSW 559
QY 586 WDKENKENVKMSAVLK-----NKETCK-----DYKFKQKIP 617
DB 560 WETNGPVWEGMLCALSVDTSLNINNVNPTHKLTGEGNNNFEKVI FGSDSSTLSKFSERP 619
QY 618 QPLRFKEGWDGFCRKEKELYSPEKVECK-KDCDEN-----TKNKCSEYK 666
DB 620 QFLRLTWGSENFCKEKKK-----YKVLAKCKDCVDGDKNGKVCVACKQCKQYH 673
QY 667 KWLIDKKSEYKQVDKYTKDKKKMYNDNIDEVQNK-EANVYLKESKSECK-----DWNF 719
DB 674 SWIGIWDNYYKQGRYTEVKKIPLYKEDKDKVNSDDARDLYLTQLQNMKCVNGTTDENC 733
QY 720 DDKIFNESPNEVDCKKCKDE 740
DB 734 EYKCMHKTSTNSDMPESLDE 754

RESULT 26
US-10-153-273-8
; Sequence 8, Application US/10153273
; Patent No. 6962987
```

```
Qy 435 -----PA-----KNDVDIAS 444
Db 341 CPKQTVEDKKBEETCTTASPVEKPEVPHVARWRTFTTPEVFKIWRGRNKTTCBIVA 400
Qy 445 QINVDLRG---FCNYSKNNKSWNCTGTFTNKFP---GTCEPPRRQTLCL----- 490
Db 401 EM-LKDKNGRTTVGECYKETYSEWTC---ESKIKMGQHGACIPPRQKLCLHYLEKIM 456
Qy 491 -----GRTYLL-----HRGHEEDYKEHLGASIEYEAQLLKYKYEKDE 528
Db 457 TTNTELKYPATKCAAEFTLLWQYKDKNGNAEDLDEKLGGIIPEDFKQMPY----- 511
Qy 529 NALCSIIONSVADLADIKSGDIIKDYKKGWENLKV--NKDKKNBESLKI-FREKW 505
Db 512 -----TFADYRDI CLGTDISSK---KDTSGVGKVRKNDIDVDFYKISNIRYKSW 559
Qy 586 WDKENKVMKSAVLK-----NKETCK-----DYDKFOKIP 617
Db 560 WETNGPVIWEGMLCALSYDTSLNWVNFETHKLTGEGNNFEKVIFGSDSTTLTKFSERP 619
Qy 618 QFLRFKFWGDDFCERKEKIYSPESFKVECK-KQCDEN-----TCRKNKSEYK 666
Db 620 QFLRWLTWEGNFCPEQKKE-----YKVLAKCKDCDVGDKGKNGKCVACKDQCKQYH 673
Qy 667 KWIDLKSEYEKQVDKTKDKKQKQYNDIDEVKQK-BANVYLKESKECK-----DVNF 719
Db 674 SWIGIWDNKKQGRYTVKPKIPLYKEDKDVKNSDDDARDYLKTLQNMKCVNGTTDENC 733
Qy 720 DDKIFNESPNEYEDMCKKDE 740
Db 734 EYKCHKTSNSTNDMPESLDE 754

RESULT 27
US-08-487-826B-16
; Sequence 16, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2182 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-487-826B-16

Query Match 5.5%; Score 353.5; DB 1; Length 2182;
Best Local Similarity 20.8%; Pred. No. 3.1e-15;
Matches 198; Conservative 133; Mismatches 346; Indels 275; Gaps 45;

Qy 379 LYKKKNSGVDDKNNFLNDLFRKNNKNDLDDFFFKNEKYEVDLDCDCRYTATIIKSF---LNG- 434
Db 43 VYEEKVNGDAKKYIEAL--KGNLTANG-----RSSETASSIETCTLVKEYYERVNGD 94
Qy 435 ----PAKNDVDIASQINVDLRGFCGNYKSNNEKSNWCTGTFTNKPGTCPEPPRRQTL- 489
Db 95 GKRPCHCRKDAKNSDVNRFSDTLGGQCTYNRKIDSQQG-----DNKV-GACAPYRRLHCD 148
Qy 490 ----LGRTVLLHRGHEEDYKEHLGASIEYEAQLLKYK-----KEKDNALCSIIIONS 538
Db 149 YNLESIDTSTTHKLLE-----VCHAAKYEGNSINTHYTHQORTNEDSASQLCVLARS 203
Qy 539 YADLADIKSGDIIKDY-----YKKGWENL-----NKVNKDKKRNESLKI-F- 581
Db 204 FADIGDIVRGKDLVLYGDNKEKEQRKKLEOKLADI FKFIHKDVMTKNGAQERYIDDAGG 263
Qy 582 ----REKWDENKENVMK-----VMSAVLKNKET-----CKDYD---KFQK 615
Db 264 DFFQLREDWMTSNRETVMKALICHAPKEANYFIKTACNVGKGTNGQCHCIGGVPTVFDY 323
Qy 616 IPOFLRFKFWGDDFCERKEKIYSPESFKVECK-----KQCDEN----- 655
Db 324 VPQYLRFWEAEADFCRKKKKL---ENLQKQCRDYEQNLCSGNGVDCTKTYIKKGLV 380
Qy 656 ----NTCRNKCSEYKQWIDLKSEYEKQVDKTKD-----KKNQMYNDIDEVKKE 702
Db 381 IGEHCTNCSVWCRMYETWIDNQKKEFLKQKRKYETETISGGSGSGSKPKRTKRAARSSSSD 440
Qy 703 ANVYLKESKECKDVNFDD-----KIFNESPNEYEDMC-----KKCDEIKYLAIEKY 749
Db 441 DNGYESKFKYKLEKVGQVDVKFLKILNK-----EGICQKQPVGNKKAQNVDFTN-KY 494
Qy 750 PKT--KHDIYDIDFTFGDTGTPISINAMINEQQSGKOTSNTGNSSET-----SDSPVS 801
Db 495 VKTFSRTEICEPCPCWGLEKGGPPWKV-----KDKTKCSAKTKTYDPKNTIDPVL 546
Qy 802 HEPESDAAINVEKLSGDESSETRGIL-----DINDPSVTNNVNEVHDASN--TQGSV 852
Db 547 YPDKSQONILKYNFCGKAGPGGGQIKKQCYVDEHRPSSKNNNVCVEGTWDFKFTQKQ 606
Qy 853 SNTS-----DITNGHSSS--LNRTTNAQDIKIGSGNE-QSD----- 887
Db 607 TVKSYNVFWDVHMDLHDSVEWKTELKSCINNNTNGNTC---RNNNCKTKDCCCFKWV 663
Qy 888 --NQENSSSHSDNSGSLT-----IGQVPSEDN--TQTYDSQNPHRDTPN 928
Db 664 EKKQEWMAIKDHFQKQTDIVQQKGLIVFSPYGLVDLVKGGNLLQNI KDVHGDTDDIKH 723
Qy 929 ALASLPDSDKINIEGFDSDSRDSENGRGDTTSTNTHDVRRTNIVSERRVNSHDFIRNGMAN 988
Db 724 IKKLLDEDAVAVVLG-----GKDNNT-----IDKLLQH 752
Qy 989 NNAHQYITQIENNGIIRGOESAGSNVYKDNPKR---SNFSSENDHKKNIQYNSRDT 1045
Db 753 EKQAEQCKQKQBECEKKAQOESGRSAETREDERTQQPADSAGEVEEEDDDDYDEDE 812
Qy 1046 KRVREETIKLSKQNKCNBNYSMEYCTYSDE-----RNSSPGPCS-----REERKKLCC 1093
Db 813 D---DDVVQEEBEGK--EEGTVTVEVTEVTEVVEETVTEQSGVKPDCDIVGKLFEDDKSL-- 865
```

```
QY 1094 QISDYC-LKYFNFSYIEYNCIKSEIKSPYKCFKSEQSSIPYPAAGGILV 1144
Db 866 --KEACGLKYGp-----GGKEKFNKCVTPSGVSTATSGKGAICV 905

RESULT 28
US-10-087-013-7
; Sequence 7, Application US/10087013
; Patent No. 6855323
; GENERAL INFORMATION:
; APPLICANT: Arthur Scherf
; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvelle
; APPLICANT: No. 6855323utaka Fujii
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF
; FILE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: PCT/US00/24195
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/152,023
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-013-7

Query Match 5.4%; Score 348; DB 2; Length 294;
Best Local Similarity 32.9%; Pred. No. 4.3e-16;
Matches 94; Conservative 37; Mismatches 87; Indels 68; Gaps 15;

QY 456 CNYKSNNEKSWNC-TGFTNKPFGTCPPRRQTCLGRTYLLHRGHEIDYK-----EHL 508
Db 13 CNAKURKKNQECQDQNTFVDCNEGVCMPRRKSIHNLTL-----EQGTNKYQLREAF 67
QY 509 LGASIEYEAQLLYKXKEDNALCSIIQN-----SYADLADIIGSDIHKDY 555
Db 68 IKCAAKETNLLWDKYK-NDKREABELLKKKIPEDFMRIMFYTFGDFRDFCLENMGRDV 126
QY 556 YGKMEENLKV-NKDKKRNEESLKI PREKWDENKENVKMSAVL-----KNKE 605
Db 127 --DKVKNKVNFNNSRGRGFK--KIDPENWNNNGPQIWNGLCALIHADTKOSIKND 182
QY 606 TKCDYDK-----FQKIPQFLRFKMGDDDFCEKREKIKYISFSEFKVEK 648
Db 183 NYK-YEYVTILAKRDSNGMTLSEFAKPKFLRFVWYDDYCKERQKYLTEVAS---TC 238
QY 649 KKCD-----CDNTCKNKCSEYKWKIDLKSEYKQVDKYTKDKNK 689
Db 239 KSIDGGQLKCDRG-CNNKCDSEYKWKIRKKGEWNLQ-DKYTKDKRE 282

RESULT 29
US-10-087-013-9
; Sequence 9, Application US/10087013
; Patent No. 6855323
; GENERAL INFORMATION:
; APPLICANT: Arthur Scherf
; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
```

```
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvelle
; APPLICANT: No. 6855323utaka Fujii
; APPLICANT: Joseph Smith
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF
; FILE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: PCT/US00/24195
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/152,023
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-013-9

Query Match 5.1%; Score 333.5; DB 2; Length 351;
Best Local Similarity 27.9%; Pred. No. 5.6e-15;
Matches 92; Conservative 46; Mismatches 131; Indels 61; Gaps 11;

QY 442 IASQINVDLGRFGCNYKSNNEKSWNCCTGTNTNKPFGTCPPRRQTCLG-RTYLLHRGH 500
Db 9 ISEQIEKNH-----NCKKTEDAKWCENTKLGDEGVCMPRRNLCHVYLTKLNDISK 64
QY 501 EEDYKEHLGLASIEYEAQLLYKXKEDNAL-----CSIIQNSYADLADIIGK 548
Db 65 EEDLEAFIKSAAAEFTLLRQYNSKNVEDDKILHRDMIPPEFSPFMYTFGDIYDCLD 124
QY 549 SDI---IKDY----YGKMBENLKNVKNKDNESLKI PREKWDENKENVK----- 595
Db 125 TDISEKIADHVTAKKKITAVFQKIGSKTNGKKVLE--REGWVKYGLSIKGMCLAL 182
QY 596 -----VMSAVLKNKETCKDYDKFQKIPQFLRFKMGDDDFCEKREKIKY 639
Db 183 SYNTETKMDGVRTYLMKYIKNDIKYELSEFASRPPFLRWVTWEGDFVGRKKSLV 242
QY 640 SFSEFKVEK-----KKDCDEN-----TCNKNKCSEYKWKIDLKSEYKQVDKYTKDN 688
Db 243 SLKKKCDSCITLRNGTNGTKTCDNENCGACTQCEKIKKMERWKKYSSQKKKFLYKN 302
QY 689 KMDYNDIDEVKNKEANYLKEKSECKDVN 718
Db 303 SATYNN--GLAVKEANSETYKNDPEVTEAN 330

RESULT 30
US-10-087-013-11
; Sequence 11, Application US/10087013
; Patent No. 6855323
; GENERAL INFORMATION:
; APPLICANT: Arthur Scherf
; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvelle
; APPLICANT: No. 6855323utaka Fujii
; APPLICANT: Joseph Smith
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF
; FILE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; CURRENT FILING DATE: 2002-02-21
```





Db 240 DCTKIYKKGKLVIGEHCTNCSVWCRMYETWIDNQKEFLKQKRKYETEISGGSGKSPK 299

Qy 690 KMYDNIDEVKNKEANVYLKSKSKCDVNFDD-----KIFNESPNEDYEDMC-----K 736

Db 300 RTKRAARSSSSDDNGYESKFYKLVKEGVQDVDFKILNK-----EGICQKQPOVGNE 354

Qy 737 KCDEIKYLNIEIKYKPT--KHDIYDIDTFSDFGDTGISINANINEQOQSGKDTSTNGSE 794

Db 355 KADNVDTNE-KYVKTFSRTEICEPCPCWGLEKGGPPWKV-----KGDTCGSAKTK 405

Qy 795 T-----SDSPVSHESPESDAINEKLSGDESSESTRGIL-----DINDPSVTNNVNE 841

Db 406 TYDPKNITDIPVLVYDKSQNLKYNKFCCKGAPGGQIKKWQCYDEHRPSSKNNNC 465

Qy 842 VHDASN--TQG 850

Db 466 VEGTWKFTQG 476

RESULT 34

US-10-153-273-10

; Sequence 10, Application US/10153273

; Patent No. 6962987

; GENERAL INFORMATION:

; APPLICANT: Sim, Kim L.

; Chitnis, Chetan

; Miller, Louis H.

; Peterson, David S.

; Su, Xin-zhaun

; Wellens, Thomas E.

; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe Martens Olson & Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: California

; COUNTRY: US

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/153,273

; FILING DATE: 21-May-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/210,288

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Fuller, Michael

; REGISTRATION NUMBER: 36,516

; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 235-8550

; TELEFAX: (619) 235-0176

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 700 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Plasmodium falciparum

; SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-10-153-273-10

Query Match 5.0%; Score 325; DB 2; Length 700;

Best Local Similarity 24.4%; Pred. No. 5.8e-14;

Matches 120; Conservative 70; Mismatches 161; Indels 140; Gaps 22;

Qy 478 GTCEPERRRLTLC-----LGRTVLLHRGHEEDYKEHLLGASIEAQLKYK-----KE 525

Db 8 GACAPYRRLHLCDYNLESIDTSTTHKLLE-----VCAAKYEGNSINTHYTQHQTNE 62

Qy 526 KBNALCSIQNSIYADLADIIGSDIHKDY-----YGKMEENL-----NKVAKDKKRNE 575

Db 63 DSASQLCTVLARSFADIGDIVRGKDLVLYGDNKEKEQRKLEQKLDIPKKIHKDVMTN 122

Qy 576 ESLKIF-----REKWDENKENVWK-----VMSAVLNKNET----- 606

Db 123 GAQERYIDDAKGGDFFQLREDMWTNSRETVMKALICHAPKEANYFIKTACNVGKGTNGQC 182

Qy 607 -CKDYD---KFQKIPQFLRFKWEKGGDFCEKREKIYSPSPKVECKK-----K 651

Db 183 HCIGGDVPTFYDVVPOYLRWFEEWAEDFCRKKKKL---ENLQKQCRDYEQNLYCSNGY 239

Qy 652 DCDE-----NTCKNKCSEYKKWIDLKSEYKQVDTKYTKD-----KNK 689

Db 240 DCTKIYKKGKLVIGESHCTNCSVWCRMYETWIDNQKEFLKQKRKYETEISGGSGKSPK 299

Qy 690 KMYDNIDEVKNKEANVYLKSKSKCDVNFDD-----KIFNESPNEDYEDMC-----K 736

Db 300 RTKRAARSSSSDDNGYESKFYKLVKEGVQDVDFKILNK-----EGICQKQPOVGNE 354

Qy 737 KCDEIKYLNIEIKYKPT--KHDIYDIDTFSDFGDTGISINANINEQOQSGKDTSTNGSE 794

Db 355 KADNVDTNE-KYVKTFSRTEICEPCPCWGLEKGGPPWKV-----KGDTCGSAKTK 405

Qy 795 T-----SDSPVSHESPESDAINEKLSGDESSESTRGIL-----DINDPSVTNNVNE 841

Db 406 TYDPKNITDIPVLVYDKSQNLKYNKFCCKGAPGGQIKKWQCYDEHRPSSKNNNC 465

Qy 842 VHDASN--TQG 850

Db 466 VEGTWKFTQG 476

RESULT 35

US-10-087-013-10

; Sequence 10, Application US/10087013

; Patent No. 6855323

; GENERAL INFORMATION:

; APPLICANT: Arthur Scherf

; APPLICANT: Louis H. Miller

; APPLICANT: Benoit Gamain

; APPLICANT: Dror I. Baruch

; APPLICANT: Pierre Buffet

; APPLICANT: Christine Scheidig

; APPLICANT: Jurg Gysin

; APPLICANT: Bruno Pouvelle

; APPLICANT: No. 6855323utaka Fujii

; APPLICANT: Joseph Smith

; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF

; PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1

; TITLE OF INVENTION: (PEMPE1) THAT MEDIATES ADHESION TO CHONDROITIN SULFATE A

; FILE REFERENCE: NIH176.001C1

; CURRENT APPLICATION NUMBER: US/10/087,013

; CURRENT FILING DATE: 2002-02-21

; PRIOR APPLICATION NUMBER: PCT/US00/24195

; PRIOR FILING DATE: 2000-09-01

; PRIOR APPLICATION NUMBER: 60/152,023

; PRIOR FILING DATE: 1999-09-01

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; TYPE: PRT

; LENGTH: 311

; ORGANISM: Plasmodium falciparum

US-10-087-013-10





## RESULT 39

US-08-487-826B-27  
; Sequence 27, Application US/08487826B  
; Patent No. 5993827  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellem, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,826B  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israel, Ned  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: NIH121.001CP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 277 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
US-08-487-826B-27

Query Match 4.7%; Score 301.5; DB 1; Length 277;  
Best Local Similarity 31.1%; Pred. No. 6.5e-13;  
Matches 64; Conservative 25; Mismatches 112; Indels 5; Gaps 3;  
  
Qy 468 CTGTTNKFPGTCPPRRQTLCLGRYLLHRGHEEDYKHLGLGASIEAQLLYKYKEKD 527  
Db 15 CXXXXXXXVCPVPPRQELCLGNIXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 74  
  
Qy 528 ENALCSIIQNSYADLADIIKGSIIIDYKYGKMEENLNKVNKKRNEESLKIFREKWD 587  
Db 75 XXEYCKIINTFADIRDIIGTGYNDLSNRXXXXXXXXXXXXXXXXXXXXKNDLFRDEWK 134  
  
Qy 588 ENKENVWKMVASLKNKCTCKDYDKFKIQQFLRWPKWGDGDDFCCKRKEKIKYSPESFKVE 647  
Db 135 VIKKDVNWNIS-WFXXXXXC-XXXXXXXIQQFFRWFSEWGDYCDQTKMI---ETLKVE 189  
  
Qy 648 CKKDCDCEKCKSEYKKWIDLK 673  
Db 190 CXXXXCXDDNCKSKNSYKEWISKK 215

## RESULT 40

US-09-210-288-15  
; Sequence 15, Application US/09210288  
; Patent No. 6392026  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellem, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/210,288  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fuller, Michael  
; REGISTRATION NUMBER: 36,516  
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 277 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
US-09-210-288-15

Query Match 4.7%; Score 301.5; DB 2; Length 277;  
Best Local Similarity 31.1%; Pred. No. 6.5e-13;  
Matches 64; Conservative 25; Mismatches 112; Indels 5; Gaps 3;  
  
Qy 468 CTGTTNKFPGTCPPRRQTLCLGRYLLHRGHEEDYKHLGLGASIEAQLLYKYKEKD 527  
Db 15 CXXXXXXXVCPVPPRQELCLGNIXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 74  
  
Qy 528 ENALCSIIQNSYADLADIIKGSIIIDYKYGKMEENLNKVNKKRNEESLKIFREKWD 587  
Db 75 XXEYCKIINTFADIRDIIGTGYNDLSNRXXXXXXXXXXXXXXXXXXXXKNDLFRDEWK 134  
  
Qy 588 ENKENVWKMVASLKNKCTCKDYDKFKIQQFLRWPKWGDGDDFCCKRKEKIKYSPESFKVE 647  
Db 135 VIKKDVNWNIS-WFXXXXXC-XXXXXXXIQQFFRWFSEWGDYCDQTKMI---ETLKVE 189  
  
Qy 648 CKKDCDCEKCKSEYKKWIDLK 673  
Db 190 CXXXXCXDDNCKSKNSYKEWISKK 215

Search completed: November 21, 2005, 20:49:10  
Job time : 75 secs

**This Page Blank (uspto)**

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:26:58 ; Search time 257 Seconds  
(without alignment)  
3321.748 Million cell updates/sec

Title: US-10-677-980-2  
Perfect score: 6481  
Sequence: 1 MKGYFNIFLYPLIFLYNVI.....VQETNISDYSEYNNKMY 1210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6481	100.0	1210	2	Q96724_PLAFA
2	6471	99.8	1210	2	Q8WS31_PLAFA
3	6471	99.8	1210	2	Q76NM5_PLAF7
4	6091	94.0	1137	2	Q8WM32_PLAFA
5	6085	93.9	1137	2	Q8MTI9_PLAFA
6	6080	93.8	1137	2	Q8MTI8_PLAFA
7	6077	93.8	1137	2	Q8MTI6_PLAFA
8	6077	93.8	1137	2	Q8MTI7_PLAFA
9	5606.5	86.5	1282	2	Q5SEX8_PLARE
10	3374	52.1	616	2	Q8MM04_PLAFA
11	3368	52.0	616	2	Q8MM45_PLAFA
12	3367	52.0	616	2	Q8T9I8_PLAFA
13	3365	51.9	616	2	Q8I049_PLAFA
14	3364	51.9	616	2	Q8MM74_PLAFA
15	3364	51.9	616	2	Q8T9N3_PLAFA
16	3363	51.9	616	2	Q8I787_PLAFA
17	3360	51.8	616	2	Q8I788_PLAFA
18	3359	51.8	616	2	Q8I786_PLAFA
19	3319	51.2	606	2	Q66PM7_PLAFA
20	3309	51.1	606	2	Q66PL8_PLAFA
21	3144	48.5	616	2	Q8I760_PLARE
22	1700	26.2	1475	2	Q25842_PLAFA
23	1596.5	24.6	1462	2	Q8IB88_PLAF7
24	1557	24.0	1435	1	EBAL_PLAFC
25	1547	23.9	1435	2	Q9NG63_PLAFA
26	1537.5	23.7	1421	2	Q05644_PLAFA
27	1524	23.5	1402	2	Q9SVT1_PLAFA
28	1514.5	23.4	1433	2	Q9N9G9_PLARE
29	1210.5	18.7	614	2	Q8I796_PLAFA
30	1209.5	18.7	616	2	Q8I791_PLAFA
31	1209.5	18.7	616	2	Q8I797_PLAFA

32	1207.5	18.6	614	2	Q8I7A0_PLAFA	Q8I7A0 plasmodium
33	1206.5	18.6	616	2	Q90585_PLAFA	Q90585 plasmodium
34	1206.5	18.6	616	2	Q25739_PLAFA	Q25739 plasmodium
35	1204.5	18.6	614	2	Q25736_PLAFA	Q25736 plasmodium
36	1204.5	18.6	614	2	Q8I790_PLAFA	Q8I790 plasmodium
37	1204.5	18.6	616	2	Q8I0C8_PLAFA	Q8I0C8 plasmodium
38	1204.5	18.6	616	2	Q27390_PLAFA	Q27390 plasmodium
39	1203.5	18.6	616	2	Q27248_PLAFA	Q27248 plasmodium
40	1202.5	18.6	616	2	Q27361_PLAFA	Q27361 plasmodium
41	1202.5	18.6	616	2	Q8I795_PLAFA	Q8I795 plasmodium
42	1199.5	18.5	616	2	Q25735_PLAFA	Q25735 plasmodium
43	1198.5	18.5	614	2	Q8I0U2_PLAFA	Q8I0U2 plasmodium
44	1198.5	18.5	616	2	Q8I792_PLAFA	Q8I792 plasmodium
45	1198.5	18.5	616	2	Q8I793_PLAFA	Q8I793 plasmodium

ALIGNMENTS

RESULT 1  
Q96724\_PLAFA  
ID Q96724\_PLAFA PRELIMINARY; PRT; 1210 AA.  
AC Q96724;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Erythrocyte binding protein BAEBL.  
GN Name=baebl;  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP NUCLEOTIDE SEQUENCE  
RC STRAIN=Indochina III/CDC;  
RX MEDLINE=21211106; PubMed=11309486; DOI=10.1073/pnas.081075398;  
RA Mayer D.C., Kaneko O., Hudson-Taylor D.E., Reid M.E., Miller L.H.;  
RT "Characterization of a Plasmodium falciparum erythrocyte-binding  
RT protein paralogous to EBA-175.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:5222-5227(2001).  
DR EMBL; AF332918; AAK49521.1; -, Genomic\_DNA.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; P:receptor activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR008602; Duffy binding.  
DR Pfam; PF05424; Duffy\_binding, 1.  
SQ SEQUENCE 1210 AA; 140527 MW; F0E56B397C8E8271 CRC64;

Query Match	100.0%;	Score 6481;	DB 2;	Length 1210;
Best Local Similarity	100.0%;	Pred. No. 6.1e-264;		
Matches 1210;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MKGYNFYFLIPLIFLYNVRINESII	GRTLYNVRQDESSDISRVNSPELNNHNTIYDS	60
Db	1	MKGYNFYFLIPLIFLYNVRINESII	GRTLYNVRQDESSDISRVNSPELNNHNTIYDS	60
Qy	61	DYEDVNNKLSNFVENKSVKKKSLSP	INNKTYSYDIIIPPSYSVRNDKNSLSNENSG	120
Db	61	DYEDVNNKLSNFVENKSVKKKSLSP	INNKTYSYDIIIPPSYSVRNDKNSLSNENSG	120
Qy	121	NTNSNFPANTSEISIGKDNKQYTFI	QKRTLFLFACGKKRSIKWICRENSSEKITVCVPDRK	180
Db	121	NTNSNFPANTSEISIGKDNKQYTFI	QKRTLFLFACGKKRSIKWICRENSSEKITVCVPDRK	180
Qy	181	IQLCVANFLNRLTMEKFKFEI	FLISVNTAKLLYNQVEKDPISIFCNELRNSFSDFRSS	240
Db	181	IQLCVANFLNRLTMEKFKFEI	FLISVNTAKLLYNQVEKDPISIFCNELRNSFSDFRSS	240
Qy	241	FIGDDMDFGNTDRVKGYINTKFS	DYKKNVEKLNNIKKEWKEKNKANLNNHMI	VNHKG 300
Db	241	FIGDDMDFGNTDRVKGYINTKFS	DYKKNVEKLNNIKKEWKEKNKANLNNHMI	VNHKG 300
Qy	301	NISKECAIIIPABEPQINLWIK	WENENFLMEKGRFLNIDKDCVENKKEACFCGCGRLPCS	360

Db 301 NISKECAIIPAEBPQINLWIKEWNEFLMEKRLFLNIKDKCVENKYYEACFGCRLPCS 360  
Qy 361 SYTSFMKSKTQMEVLTLNLYKQSGVDKNNFLNDLFKQNNKNDLDDFFKNEKEYDDLCD 420  
Db 361 SYTSFMKSKTQMEVLTLNLYKQSGVDKNNFLNDLFKQNNKNDLDDFFKNEKEYDDLCD 420  
Qy 421 CRYTATIIKSLFNGPAKNDVDIASQINVDLRFKQNGYKSNNEKSWNCTGTFTNKPFGTC 480  
Db 421 CRYTATIIKSLFNGPAKNDVDIASQINVDLRFKQNGYKSNNEKSWNCTGTFTNKPFGTC 480  
Qy 481 EPPRQTLCLGRYLLHRGHEEDYKEHLLGASIEAQLLYKKEKDNALCSIIQNSYA 540  
Db 481 EPPRQTLCLGRYLLHRGHEEDYKEHLLGASIEAQLLYKKEKDNALCSIIQNSYA 540  
Qy 541 DLADIIKGSIIKDYKQKMEENLVNKKVKKRNEESLKIIFREKWDENKENVKWSAV 600  
Db 541 DLADIIKGSIIKDYKQKMEENLVNKKVKKRNEESLKIIFREKWDENKENVKWSAV 600  
Qy 601 LKNKETCKDYDKFKOKIPOFLRWFKEWGDDECEKKEKIYSFESFKVECKKDCDENTCKN 660  
Db 601 LKNKETCKDYDKFKOKIPOFLRWFKEWGDDECEKKEKIYSFESFKVECKKDCDENTCKN 660  
Qy 661 KCSEYKKWIDLKSEYKQVDKYTKDKNKKMYDNIDEVKKEANVYLKEKSKCKDVNFD 720  
Db 661 KCSEYKKWIDLKSEYKQVDKYTKDKNKKMYDNIDEVKKEANVYLKEKSKCKDVNFD 720  
Qy 721 DKIFNESPNEYEDMCKCDEIKYLNIEIKYPTKHDYIDITFSDTFDGGTPIISINANINE 780  
Db 721 DKIFNESPNEYEDMCKCDEIKYLNIEIKYPTKHDYIDITFSDTFDGGTPIISINANINE 780  
Qy 781 QQSGKDTNTGNSSETSDSPVSHPEPESDAAINVEKLSGDESSSTRGLDINDSVTNVN 840  
Db 781 QQSGKDTNTGNSSETSDSPVSHPEPESDAAINVEKLSGDESSSTRGLDINDSVTNVN 840  
Qy 841 EVHDSNTQGSVNTSDITNGHSSSLNRTTNAQDIKIGSGNQSDNQNSHSSNSG 900  
Db 841 EVHDSNTQGSVNTSDITNGHSSSLNRTTNAQDIKIGSGNQSDNQNSHSSNSG 900  
Qy 901 SLTIGQVPESENTQNTYDSQNPHRDTPNALASPSDDKINEIEGFSRSDSENGRGTTS 960  
Db 901 SLTIGQVPESENTQNTYDSQNPHRDTPNALASPSDDKINEIEGFSRSDSENGRGTTS 960  
Qy 961 NTHDVRNTIVSERVNSHDFIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020  
Db 961 NTHDVRNTIVSERVNSHDFIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020  
Qy 1021 NPKESNFSSENHKKNIQEVNSRDTKVRBEIILKSKQNCNNEYSMEYCTYSDERNSSP 1080  
Db 1021 NPKESNFSSENHKKNIQEVNSRDTKVRBEIILKSKQNCNNEYSMEYCTYSDERNSSP 1080  
Qy 1081 GPCSRBRKLCQISDYCLKFNFYSIYNYCTIKSEIKSPYKCPKSEGOSSIPYPAAG 1140  
Db 1081 GPCSRBRKLCQISDYCLKFNFYSIYNYCTIKSEIKSPYKCPKSEGOSSIPYPAAG 1140  
Qy 1141 GILVIVILLSSARMKSENEVDIGESNIEATEPEENNYLNKLSRIENQVQETNIDSYS 1200  
Db 1141 GILVIVILLSSARMKSENEVDIGESNIEATEPEENNYLNKLSRIENQVQETNIDSYS 1200  
Qy 1201 EYNYNEKNMY 1210  
Db 1201 EYNYNEKNMY 1210

## RESULT 2

Q8WS31.PLAFA  
ID Q8WS31.PLAFA PRELIMINARY; PRT; 1210 AA.  
AC Q8WS31;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
DE Erythrocyte binding antigen 140.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=5833;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=3D7;  
RA Thompson J.K., Triglia T., Reed M.B., Cowman A.F.;  
RT "A novel ligand from Plasmodium falciparum that binds to a sialic  
RT acid-containing receptor on the surface of human erythrocytes.";  
RL Mol. Microbiol. 0:0-0(2002).  
DR EMBL; AF384554; AAL58319.1; -; Genomic DNA.  
DR GO:0016021; C:integral to membrane; IEA.  
DR GO:0004872; F:receptor activity; IEA.  
DR GO:0009405; P:pathogenesis; IEA.  
DR InterPro: IPR008602; Duff binding; IEA.  
DR Pfam: PF05424; Duff binding; 1.  
SQ SEQUENCE 1210 AA; 140595 MW; PF44FBB8DA69D605 CRC64;

Query Match 99.8%; Score 6471; DB 2; Length 1210;

Best Local Similarity 99.8%; Pred. No. 1.6e-263;

Matches 1207; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKGYFNIYFLIPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELANNHKNTIYDS 60

Db 1 MKGYFNIYFLIPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELANNHKNTIYDS 60

Qy 61 DYEDVNNKLINSFVENKSVKKGSLSFNNKTKSYDIIPPSYRNDKFNLSLENEDNSG 120

Db 61 DYEDVNNKLINSFVENKSVKKGSLSFNNKTKSYDIIPPSYRNDKFNLSLENEDNSG 120

Qy 121 NTNSNPFANTSEISGKDNKQYTFIOKRTHLFACGIRKRSIKWICRENSSEKITVCPVDRK 180

Db 121 NTNSNPFANTSEISGKDNKQYTFIOKRTHLFACGIRKRSIKWICRENSSEKITVCPVDRK 180

Qy 181 IQLCVAFNLSRLTMEKFEKFLISVNTAEKLLYNKNEGKOPSIFCNELRNSFSFPRS 240

Db 181 IQLCVAFNLSRLTMEKFEKFLISVNTAEKLLYNKNEGKOPSIFCNELRNSFSFPRS 240

Qy 241 FIGDDMDFGNTDRVKGYNTRKFSDDYKQVKEKLANIKKEWKEKKNANLWHEMIVNHKG 300

Db 241 FIGDDMDFGNTDRVKGYNTRKFSDDYKQVKEKLANIKKEWKEKKNANLWHEMIVNHKG 300

Qy 301 NISKECAIIPAEBPQINLWIKEWNEFLMEKRLFLNIKDKCVENKYYEACFGCRLPCS 360

Db 301 NISKECAIIPAEBPQINLWIKEWNEFLMEKRLFLNIKDKCVENKYYEACFGCRLPCS 360

Qy 361 SYTSFMKSKTQMEVLTLNLYKQNGVDKNNFLNDLFKQNNKNDLDDFFKNEKEYDDLCD 420

Db 361 SYTSFMKSKTQMEVLTLNLYKQNGVDKNNFLNDLFKQNNKNDLDDFFKNEKEYDDLCD 420

Qy 421 CRYTATIIKSLFNGPAKNDVDIASQINVDLRFKQNGYKSNNEKSWNCTGTFTNKPFGTC 480

Db 421 CRYTATIIKSLFNGPAKNDVDIASQINVDLRFKQNGYKSNNEKSWNCTGTFTNKPFGTC 480

Qy 481 EPPRQTLCLGRYLLHRGHEEDYKEHLLGASIEAQLLYKKEKDNALCSIIQNSYA 540

Db 481 EPPRQTLCLGRYLLHRGHEEDYKEHLLGASIEAQLLYKKEKDNALCSIIQNSYA 540

Qy 541 DLADIIKGSIIKDYKQKMEENLVNKKVKKRNEESLKIIFREKWDENKENVKWSAV 600

Db 541 DLADIIKGSIIKDYKQKMEENLVNKKVKKRNEESLKIIFREKWDENKENVKWSAV 600

Qy 601 LKNKETCKDYDKFKOKIPOFLRWFKEWGDDECEKKEKIYSFESFKVECKKDCDENTCKN 660

Db 601 LKNKETCKDYDKFKOKIPOFLRWFKEWGDDECEKKEKIYSFESFKVECKKDCDENTCKN 660

Qy 661 KCSEYKKWIDLKSEYKQVDKYTKDKNKKMYDNIDEVKKEANVYLKEKSKCKDVNFD 720

Db 661 KCSEYKKWIDLKSEYKQVDKYTKDKNKKMYDNIDEVKKEANVYLKEKSKCKDVNFD 720

Qy 721 DKIFNESPNEYEDMCKCDEIKYLNIEIKYPTKHDYIDITFSDTFDGGTPIISINANINE 780

Db 721 DKIFNESPNEYEDMCKCDEIKYLNIEIKYPTKHDYIDITFSDTFDGGTPIISINANINE 780

Qy 781 QQSGKDTNTGNSSETSDSPVSHPEPESDAAINVEKLSGDESSSTRGLDINDSVTNVN 840

```
Db 781 QSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSESTRGILDINDPSVTNNVN 840
Qy 841 EVHDASNTQGSVNTSDITNGHSSSLNRTTNAODIKIGRSGNQSDNQENSSHSSDMSG 900
Db 841 EVHDASNTQGSVNTSDITNGHSSSLNRTTNAODIKIGRSGNQSDNQENSSHSSDMSG 900
Qy 901 SLTIGQVPSDNTQNTYDSQNPDRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960
Db 901 SLTIGQVPSDNTQNTYDSQNPDRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960
Qy 961 NTHDVRRTNIVSERVNSHDFIRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Db 961 NTHDVRRTNIVSERVNSHDFIRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Qy 1021 NPKRSNFSSENDHKNTQYNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDERNSSP 1080
Db 1021 NPKRSNFSSENDHKNTQYNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDERNSSP 1080
Qy 1081 GPCSRERKKLCCQISDYCLAYFNFYSTIYNNCIKSEIKSPYKCFKSEGGSSIPYFAAG 1140
Db 1081 GPCSRERKKLCCQISDYCLAYFNFYSTIYNNCIKSEIKSPYKCFKSEGGSSIPYFAAG 1140
Qy 1141 GILVIVILLSSASRMGKSNEEYDIGESNIEATFEENNYLNKLSRIFNQEOETNISDYS 1200
Db 1141 GILVIVILLSSASRMGKSNEEYDIGESNIEATFEENNYLNKLSRIFNQEOETNISDYS 1200
Qy 1201 EYNTNEKMY 1210
Db 1201 EYNTNEKMY 1210

RESULT 3
ID Q76NM5_PLAF7 PRELIMINARY; PRT; 1210 AA.
AC Q76NM5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Erythrocyte binding antigen 140.
GN Name=WAL13P1.60;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3D7;
RA Harris B.; Lennard N.; Clark L.; Line A.; Barron A.; Corton C.;
RA Berriman M.; Pain A.; Hall N.; Atkin R.; Chillingworth C.; Doggett J.;
RA Ormond D.; Sanders M.; Hayes R.; Hall S.; Quail M.; Barrall B.;
RL Submitted (SRP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; CAD52266.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 1210 AA; 140595 MW; FF44FB8DA69D605 CRC64;

Query Match 99.8%; Score 6471; DB 2; Length 1210;
Best Local Similarity 99.8%; Pred. No. 1.6e-263;
Matches 1207; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKGFNYFLPLFLPLFLYNNVIRINESIIGRTLYNQDESSDLSRVNSPELNHNHTNIYDS 60
Db 1 MKGFNYFLPLFLPLFLYNNVIRINESIIGRTLYNQDESSDLSRVNSPELNHNHTNIYDS 60
Qy 61 DYEDVNNKLSINSFVENKSVKKRSISLSPINNTKSYDIIIPPSYSVRNDKFNLSNEONSG 120
Db 61 DYEDVNNKLSINSFVENKSVKKRSLSFINNTKSYDIIIPPSYSVRNDKFNLSNEONSG 120
Qy 121 NTNSNNFANTSEISIGKDNKQYTFIQKRTLHFACGIKRKSIIKWICRENSSEKITVCVPRK 180
Db 121 NTNSNNFANTSEISIGKDNKQYTFIQKRTLHFACGIKRKSIIKWICRENSSEKITVCVPRK 180
```

```
Db 121 NTNSNNFANTSEISIGKDNKQYTFIQKRTLHFACGIKRKSIIKWICRENSSEKITVCVPRK 180
Qy 181 IQLCVANFLNRLTMEKFEKFIPLISVNTAEKLYNNKNEGDPISFCNELANSFSDPRS 240
Db 181 IQLCVANFLNRLTMEKFEKFIPLISVNTAEKLYNNKNEGDPISFCNELANSFSDPRS 240
Qy 241 FIGDMDFGGNTDRVKGYINTKFSDDYKERVVEKLANNIKKEWEKQKLANLWNHMI VNHKG 300
Db 241 FIGDMDFGGNTDRVKGYINKKFSDDYKERVVEKLANNIKKEWEKQKLANLWNHMI VNHKG 300
Qy 301 NISKECAIIPAEERPOINLWKWENEFMEKRIPLANKDKCVENKKEVYACFGGCRLP 360
Db 301 NISKECAIIPAEERPOINLWKWENEFMEKRIPLANKDKCVENKKEVYACFGGCRLP 360
Qy 361 SYTSFMKSKTQMEVLNLYKKNSGVYDKNFLNDLPKNNKNDLDDPFFKNEKEYDDLCD 420
Db 361 SYTSFMKSKTQMEVLNLYKKNSGVYDKNFLNDLPKNNKNDLDDPFFKNEKEYDDLCD 420
Qy 421 CRYTATIIKSGFLNGPAKNDVDIASQINVDLARGFGCNYKSNNEKSWNCTGTTNKF 480
Db 421 CRYTATIIKSGFLNGPAKNDVDIASQINVDLARGFGCNYKSNNEKSWNCTGTTNKF 480
Qy 481 EPPRQTLCLGRTYLLHRGHEEDYKEHLGASIIYEAQLLYKKEKDNALCSIIQNSYA 540
Db 481 EPPRQTLCLGRTYLLHRGHEEDYKEHLGASIIYEAQLLYKKEKDNALCSIIQNSYA 540
Qy 541 DLADIIGSDIIDYGYGKMEENLNKYNKDKKRNEESLKI PREKQWDEKENVKMWASV 600
Db 541 DLADIIGSDIIDYGYGKMEENLNKYNKDKKRNEESLKI PREKQWDEKENVKMWASV 600
Qy 601 LKNETCKDYDFOKIPOFLRWFKEWGDGDFCEKKEKIIYSPESPKVECKKDCDENTCKN 660
Db 601 LKNETCKDYDFOKIPOFLRWFKEWGDGDFCEKKEKIIYSPESPKVECKKDCDENTCKN 660
Qy 661 KCSEYKWIIDLKSEYEKQVDKYTKDKNKQWYDNIDEVKNEANVYLKEKSECKQVNF 720
Db 661 KCSEYKWIIDLKSEYEKQVDKYTKDKNKQWYDNIDEVKNEANVYLKEKSECKQVNF 720
Qy 721 DKIFNESPNEVEDMCKKDEIKYLNBIKYPKTHDIYDIDTFSDTFGDTGTPISINANE 780
Db 721 DKIFNESPNEVEDMCKKDEIKYLNBIKYPKTHDIYDIDTFSDTFGDTGTPISINANE 780
Qy 781 QQSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSESTRGILDINDPSVTNNVN 840
Db 781 QQSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSESTRGILDINDPSVTNNVN 840
Qy 841 EVHDASNTQGSVNTSDITNGHSSSLNRTTNAODIKIGRSGNQSDNQENSSHSSDMSG 900
Db 841 EVHDASNTQGSVNTSDITNGHSSSLNRTTNAODIKIGRSGNQSDNQENSSHSSDMSG 900
Qy 901 SLTIGQVPSDNTQNTYDSQNPDRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960
Db 901 SLTIGQVPSDNTQNTYDSQNPDRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960
Qy 961 NTHDVRRTNIVSERVNSHDFIRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Db 961 NTHDVRRTNIVSERVNSHDFIRNGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
Qy 1021 NPKRSNFSSENDHKNTQYNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDERNSSP 1080
Db 1021 NPKRSNFSSENDHKNTQYNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDERNSSP 1080
Qy 1081 GPCSRERKKLCCQISDYCLAYFNFYSTIYNNCIKSEIKSPYKCFKSEGGSSIPYFAAG 1140
Db 1081 GPCSRERKKLCCQISDYCLAYFNFYSTIYNNCIKSEIKSPYKCFKSEGGSSIPYFAAG 1140
Qy 1141 GILVIVILLSSASRMGKSNEEYDIGESNIEATFEENNYLNKLSRIFNQEOETNISDYS 1200
Db 1141 GILVIVILLSSASRMGKSNEEYDIGESNIEATFEENNYLNKLSRIFNQEOETNISDYS 1200
Qy 1201 EYNTNEKMY 1210
Db 1201 EYNTNEKMY 1210
```

RESULT 4									
Q8MM32	PLAFA	PRELIMINARY;	PRT;	1137	AA.				
ID	Q8MT19	PLAFA	PRELIMINARY;	PRT;	1137	AA.			
AC	Q8MT19	PLAFA	PRELIMINARY;	PRT;	1137	AA.			
DT	01-OCT-2002	(Tremblrel. 22, Created)							
DT	01-OCT-2002	(Tremblrel. 22, Last sequence update)							
DT	01-FEB-2005	(Tremblrel. 29, Last annotation update)							
DE	Erythrocyte binding protein.								
OS	Plasmodium falciparum.								
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.								
OX	NCBI_TaxID=5833;								
RN	NUCLEOTIDE SEQUENCE.								
RA	Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;								
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.								
DR	EMBL; AF506832; AAM28887.1; -; Genomic DNA.								
DR	EMBL; AF506834; AAM28889.1; -; Genomic DNA.								
DR	EMBL; AY099885; AAM51621.1; -; Genomic DNA.								
DR	EMBL; AF506833; AAM28888.1; -; Genomic DNA.								
DR	GO; GO:0016021; C:integral to membrane; IEA.								
DR	GO; GO:0004872; F:receptor activity; IEA.								
DR	GO; GO:0009405; P:pathogenesis; IEA.								
DR	InterPro; IPR008602; Duffy binding.								
DR	Pfam; PF05424; Duffy binding; 1.								
SQ	SEQUENCE 1137 AA; 132218 MW; 8E1DC6CFDDA6083 CRC64;								
Query Match 94.0%; Score 6091; DB 2; Length 1137;									
Best Local Similarity 99.9%; Pred. No. 1.3e-247;									
Matches 1133; Conservative 1; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	MKGYNFYFLPLFLFLYLVIRINESIIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS	60						
Db	1	MKGYNFYFLPLFLFLYLVIRINESIIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS	60						
Qy	61	DYEDVNNKLINSFVENKSVKKRSLSPINNKTKSYDIIIPPSYSYRNDKFNLSLENEDNSG	120						
Db	61	DYEDVNNKLINSFVENKSVKKRSLSPINNKTKSYDIIIPPSYSYRNDKFNLSLENEDNSG	120						
Qy	121	NTNSNFPANTSEISIGKDNKQYTFIQKRTLFCAGIKRKSIIKMICRNSSEKITVCVPDRK	180						
Db	121	NTNSNFPANTSEISIGKDNKQYTFIQKRTLFCAGIKRKSIIKMICRNSSEKITVCVPDRK	180						
Qy	181	IQLCVANFLNSRLTMEKPKFEIPLISVNTAKLLYNKNEGKPSIFCNELRNSFSDPRSS	240						
Db	181	IQLCVANFLNSRLTMEKPKFEIPLISVNTAKLLYNKNEGKPSIFCNELRNSFSDPRSS	240						
Qy	241	FIGDDMPGGNTDRVKGYNITKFSDDYKKNVEKLANIKKEWKEKKNANLWNHMI VNHKG	300						
Db	241	FIGDDMPGGNTDRVKGYNITKFSDDYKKNVEKLANIKKEWKEKKNANLWNHMI VNHKG	300						
Qy	301	NISKECAIIIPAEBPQIINLWKENNFLEMEKRLFLNI KDKCVENKYEACFGGCRLPSCS	360						
Db	301	NISKECAIIIPAEBPQIINLWKENNFLEMEKRLFLNI KDKCVENKYEACFGGCRLPSCS	360						
Qy	361	SYTSFMKSKTQMEVLNLYKKNSGYDKNLFNDLFQKNKNDLDDFFKNEKEYDDLCD	420						
Db	361	SYTSFMKSKTQMEVLNLYKKNSGYDKNLFNDLFQKNKNDLDDFFKNEKEYDDLCD	420						
Qy	421	CRYTATIIKSLFNGPAKNDVDIASQINVDNLRGFCYKGNKNSWNCTGTFNKPFGTC	480						
Db	421	CRYTATIIKSLFNGPAKNDVDIASQINVDNLRGFCYKGNKNSWNCTGTFNKPFGTC	480						
Qy	481	EPFRQTLCLGRTYLLHRGHEEDYKHLGLGASIIYEAQLLYKKEKDNALCSIIQNSYA	540						
Db	481	EPFRQTLCLGRTYLLHRGHEEDYKHLGLGASIIYEAQLLYKKEKDNALCSIIQNSYA	540						
Qy	541	DLADIIKGSIIIDKYGKMEENLKNYKDKGRNEESLKIIFREKWDENKENYKVMASV	600						
Db	541	DLADIIKGSIIIDKYGKMEENLKNYKDKGRNEESLKIIFREKWDENKENYKVMASV	600						
Qy	601	LKNKETCKDYKFKQIKIPLQFLRWPKWGGDDFCCKRKEKISFSEKCKKDCDENTCKN	660						
RESULT 5									
Q8MT19	PLAFA	PRELIMINARY;	PRT;	1137	AA.				
ID	Q8MT19	PLAFA	PRELIMINARY;	PRT;	1137	AA.			
AC	Q8MT19	PLAFA	PRELIMINARY;	PRT;	1137	AA.			
DT	01-OCT-2002	(Tremblrel. 22, Created)							
DT	01-OCT-2002	(Tremblrel. 22, Last sequence update)							
DT	01-OCT-2003	(Tremblrel. 25, Last annotation update)							
DE	Erythrocyte-binding protein.								
OS	Plasmodium falciparum.								
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.								
OX	NCBI_TaxID=5833;								
RN	NUCLEOTIDE SEQUENCE.								
RA	Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;								
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.								
DR	EMBL; AY099886; AAM51622.1; -; Genomic DNA.								
DR	GO; GO:0016021; C:integral to membrane; IEA.								
DR	GO; GO:0004872; F:receptor activity; IEA.								
DR	GO; GO:0009405; P:pathogenesis; IEA.								
DR	InterPro; IPR008602; Duffy binding.								
DR	Pfam; PF05424; Duffy binding; 1.								
SQ	SEQUENCE 1137 AA; 132245 MW; 2D026021BC5A606D CRC64;								
Query Match 93.9%; Score 6085; DB 2; Length 1137;									
Best Local Similarity 99.8%; Pred. No. 2.4e-247;									
Matches 1132; Conservative 1; Mismatches 1; Indels 0; Gaps 0;									
Qy	1	MKGYNFYFLPLFLFLYLVIRINESIIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS	60						
Db	1	MKGYNFYFLPLFLFLYLVIRINESIIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS	60						
Qy	61	DYEDVNNKLINSFVENKSVKKRSLSPINNKTKSYDIIIPPSYSYRNDKFNLSLENEDNSG	120						
Db	61	DYEDVNNKLINSFVENKSVKKRSLSPINNKTKSYDIIIPPSYSYRNDKFNLSLENEDNSG	120						
Qy	121	NTNSNFPANTSEISIGKDNKQYTFIQKRTLFCAGIKRKSIIKMICRNSSEKITVCVPDRK	180						
Db	121	NTNSNFPANTSEISIGKDNKQYTFIQKRTLFCAGIKRKSIIKMICRNSSEKITVCVPDRK	180						
Qy	181	IQLCVANFLNSRLTMEKPKFEIPLISVNTAKLLYNKNEGKPSIFCNELRNSFSDPRSS	240						
Db	181	IQLCVANFLNSRLTMEKPKFEIPLISVNTAKLLYNKNEGKPSIFCNELRNSFSDPRSS	240						
Qy	241	FIGDDMPGGNTDRVKGYNITKFSDDYKKNVEKLANIKKEWKEKKNANLWNHMI VNHKG	300						
Db	241	FIGDDMPGGNTDRVKGYNITKFSDDYKKNVEKLANIKKEWKEKKNANLWNHMI VNHKG	300						
Qy	301	NISKECAIIIPAEBPQIINLWKENNFLEMEKRLFLNI KDKCVENKYEACFGGCRLPSCS	360						
Db	301	NISKECAIIIPAEBPQIINLWKENNFLEMEKRLFLNI KDKCVENKYEACFGGCRLPSCS	360						
Qy	361	SYTSFMKSKTQMEVLNLYKKNSGYDKNLFNDLFQKNKNDLDDFFKNEKEYDDLCD	420						
Db	361	SYTSFMKSKTQMEVLNLYKKNSGYDKNLFNDLFQKNKNDLDDFFKNEKEYDDLCD	420						
Qy	421	CRYTATIIKSLFNGPAKNDVDIASQINVDNLRGFCYKGNKNSWNCTGTFNKPFGTC	480						
Db	421	CRYTATIIKSLFNGPAKNDVDIASQINVDNLRGFCYKGNKNSWNCTGTFNKPFGTC	480						
Qy	481	EPFRQTLCLGRTYLLHRGHEEDYKHLGLGASIIYEAQLLYKKEKDNALCSIIQNSYA	540						
Db	481	EPFRQTLCLGRTYLLHRGHEEDYKHLGLGASIIYEAQLLYKKEKDNALCSIIQNSYA	540						
Qy	541	DLADIIKGSIIIDKYGKMEENLKNYKDKGRNEESLKIIFREKWDENKENYKVMASV	600						
Db	541	DLADIIKGSIIIDKYGKMEENLKNYKDKGRNEESLKIIFREKWDENKENYKVMASV	600						
Qy	601	LKNKETCKDYKFKQIKIPLQFLRWPKWGGDDFCCKRKEKISFSEKCKKDCDENTCKN	660						

Db 121 NTNSNFPANTSEISIGKDNKQYTFIQKRTHLFACGIIKRSIKWICRENSSEKITVCPDRK 180  
Qy 181 IQLCVANFLNRLTMEKFEIIFLISVNTAEKLYNKGEGDPSIFCNELRNSFSDFRSS 240  
Db 181 IQLCVANFLNRLTMEKFEIIFLISVNTAEKLYNKGEGDPSIFCNELRNSFSDFRSS 240  
Qy 241 FIGDDMDFGGTDRVKGINTKFSYDYKKEKVEKLNKIKKEWKEKKNANLWNHMI VNHKG 300  
Db 241 FIGDDMDFGGTDRVKGINTKFSYDYKKEKVEKLNKIKKEWKEKKNANLWNHMI VNHKG 300  
Qy 301 NISKECAIIPAEFPQINLWIKENWENFLMEKRLFLNFKKCVNKKYKKEACFGGCRPLCS 360  
Db 301 NISKECAIIPAEFPQINLWIKENWENFLMEKRLFLNFKKCVNKKYKKEACFGGCRPLCS 360  
Qy 361 SYTSFMKSKTQMEVLTNLYKKKNSGVVDKNNFLNDFLKKNNKNDLDDFFKNEKEYDDLCD 420  
Db 361 SYTSFMKSKTQMEVLTNLYKKKNSGVVDKNNFLNDFLKKNNKNDLDDFFKNEKEYDDLCD 420  
Qy 421 CRYTATIIKSPFLNGPAKNDVDIASQINVDLGRFGCNYSKNEKSWNCTGFTTNKFPGTC 480  
Db 421 CRYTATIIKSPFLNGPAKNDVDIASQINVDLGRFGCNYSKNEKSWNCTGFTTNKFPGTC 480  
Qy 481 EPPRQTLCLGRTYLLHRGHEEDYKEHLLGASIIYEAQLLKYKKEKDNALCSIIQNSYA 540  
Db 481 EPPRQTLCLGRTYLLHRGHEEDYKEHLLGASIIYEAQLLKYKKEKDNALCSIIQNSYA 540  
Qy 541 DLADIIKGSDDIIKDYKKGKEENLKNVNDKKNRNEESLKIIFREKWDENKENVKWSAV 600  
Db 541 DLADIIKGSDDIIKDYKKGKEENLKNVNDKKNRNEESLKIIFREKWDENKENVKWSAV 600  
Qy 601 LKNETCKDYDFQKIPOFLWFWKGGDDFCERKEKIIYSPESFKVCKKCCDENTCKN 660  
Db 601 LKNETCKDYDFQKIPOFLWFWKGGDDFCERKEKIIYSPESFKVCKKCCDENTCKN 660  
Qy 661 KCSYKWKWIDLKSEYEKQVDTYKDKNKKQYDNIDEVQKNEANVYLKESKECKOVNFD 720  
Db 661 KCSYKWKWIDLKSEYEKQVDTYKDKNKKQYDNIDEVQKNEANVYLKESKECKOVNFD 720  
Qy 721 DKIFNESPNEVEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDTFGDTGPISINANINE 780  
Db 721 DKIFNESPNEVEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDTFGDTGPISINANINE 780  
Qy 781 QOSGKDSNTGNSSETSDSPVSHEPESDAAINVEKLSGDESSESTRGILDINDPSVTNNVN 840  
Db 781 QOSGKDSNTGNSSETSDSPVSHEPESDAAINVEKLSGDESSESTRGILDINDPSVTNNVN 840  
Qy 841 EVHDSNTQGSVNTSDITNGHSSSLNRTTNAQDIKIGRSGNQSDNQNSHSSDNGS 900  
Db 841 EVHDSNTQGSVNTSDITNGHSSSLNRTTNAQDIKIGRSGNQSDNQNSHSSDNGS 900  
Qy 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960  
Db 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960  
Qy 961 NTHDVRRTNIVSERVNSHPDIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020  
Db 961 NTHDVRRTNIVSERVNSHPDIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020  
Qy 1021 NPKKSNFSSNDHKNIQEVNSRDTKRVREIILKSKONKCNNSYMEYCTYSDERNSSP 1080  
Db 1021 NPKKSNFSSNDHKNIQEVNSRDTKRVREIILKSKONKCNNSYMEYCTYSDERNSSP 1080  
Qy 1081 GPCSRERKKLCCQISDYCLKYFNPYSLEYVNCIKSBKSPYKCFKSEGOSSI 1134  
Db 1081 GPCSRERKKLCCQISDYCLKYFNPYSLEYVNCIKSBKSPYKCFKSEGOSSM 1134

## RESULT 6

Q8MT18\_PLAFA

ID Q8MT18\_PLAFA PRELIMINARY; PRT; 1137 AA.

AC Q8MT18;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Erythrocyte-binding protein.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Mayer D.G., Mu J.-B., Peng X., Su X.-Z., Miller L.H.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY099887; AAM51623.1; -; Genomic DNA.  
DR GO: 0016021; C: integral to membrane; IEA.  
DR GO: 0004872; F: receptor activity; IEA.  
DR GO: 0009405; P: pathogenesis; IEA.  
DR InterPro: IPR008602; Duff binding; IEA.  
DR Pfam: PF05424; Duff binding; 1.  
SQ SEQUENCE 1137 AA; 132293 MW; 43BDFAB02B0861D2 CRC64;

Query Match 93.8%; Score 6080; DB 2; Length 1137;  
Best Local Similarity 99.6%; Pred. No. 3.8e-247;  
Matches 1130; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKGYENIYFLIPLIFLYNVIRINESIIIGRTLYNRQDESSDISRVNSPELANNHKTNIYDS 60  
Db 1 MKGYENIYFLIPLIFLYNVIRINESIIIGRTLYNRQDESSDISRVNSPELANNHKTNIYDS 60  
Qy 61 DYEDVNNKLNINSFVENKSVKKRSLSPINNKTSYDIIIPPSYSYRNDKFNLSSENDNSG 120  
Db 61 DYEDVNNKLNINSFVENKSVKKRSLSPINNKTSYDIIIPPSYSYRNDKFNLSSENDNSG 120  
Qy 121 NTNSNFPANTSEISIGKDNKQYTFIQKRTHLFACGIIKRSIKWICRENSSEKITVCPDRK 180  
Db 121 NTNSNFPANTSEISIGKDNKQYTFIQKRTHLFACGIIKRSIKWICRENSSEKITVCPDRK 180  
Qy 181 IQLCVANFLNRLTMEKFEIIFLISVNTAEKLYNKGEGDPSIFCNELRNSFSDFRSS 240  
Db 181 IQLCVANFLNRLTMEKFEIIFLISVNTAEKLYNKGEGDPSIFCNELRNSFSDFRSS 240  
Qy 241 FIGDDMDFGGTDRVKGINTKFSYDYKKEKVEKLNKIKKEWKEKKNANLWNHMI VNHKG 300  
Db 241 FIGDDMDFGGTDRVKGINTKFSYDYKKEKVEKLNKIKKEWKEKKNANLWNHMI VNHKG 300  
Qy 301 NISKECAIIPAEFPQINLWIKENWENFLMEKRLFLNFKKCVNKKYKKEACFGGCRPLCS 360  
Db 301 NISKECAIIPAEFPQINLWIKENWENFLMEKRLFLNFKKCVNKKYKKEACFGGCRPLCS 360  
Qy 361 SYTSFMKSKTQMEVLTNLYKKKNSGVVDKNNFLNDFLKKNNKNDLDDFFKNEKEYDDLCD 420  
Db 361 SYTSFMKSKTQMEVLTNLYKKKNSGVVDKNNFLNDFLKKNNKNDLDDFFKNEKEYDDLCD 420  
Qy 421 CRYTATIIKSPFLNGPAKNDVDIASQINVDLGRFGCNYSKNEKSWNCTGFTTNKFPGTC 480  
Db 421 CRYTATIIKSPFLNGPAKNDVDIASQINVDLGRFGCNYSKNEKSWNCTGFTTNKFPGTC 480  
Qy 481 EPPRQTLCLGRTYLLHRGHEEDYKEHLLGASIIYEAQLLKYKKEKDNALCSIIQNSYA 540  
Db 481 EPPRQTLCLGRTYLLHRGHEEDYKEHLLGASIIYEAQLLKYKKEKDNALCSIIQNSYA 540  
Qy 541 DLADIIKGSDDIIKDYKKGKEENLKNVNDKKNRNEESLKIIFREKWDENKENVKWSAV 600  
Db 541 DLADIIKGSDDIIKDYKKGKEENLKNVNDKKNRNEESLKIIFREKWDENKENVKWSAV 600  
Qy 601 LKNETCKDYDFQKIPOFLWFWKGGDDFCERKEKIIYSPESFKVCKKCCDENTCKN 660  
Db 601 LKNETCKDYDFQKIPOFLWFWKGGDDFCERKEKIIYSPESFKVCKKCCDENTCKN 660  
Qy 661 KCSYKWKWIDLKSEYEKQVDTYKDKNKKQYDNIDEVQKNEANVYLKESKECKOVNFD 720  
Db 661 KCSYKWKWIDLKSEYEKQVDTYKDKNKKQYDNIDEVQKNEANVYLKESKECKOVNFD 720  
Qy 721 DKIFNESPNEVEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDTFGDTGPISINANINE 780  
Db 721 DKIFNESPNEVEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDTFGDTGPISINANINE 780

```
QY 781 QQSGKDTNTGNSSETSDPSVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 840
DB 781 QQSGKDTNTGNSSETSDPSVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 840
QY 841 EVHDASTNTGNSSETSDPSVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 900
DB 841 EVHDASTNTGNSSETSDPSVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 900
QY 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGFDSSRDSSENGRGTTS 960
DB 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGFDSSRDSSENGRGTTS 960
QY 961 NTHDVRTNIVSERRVNSHDFIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
DB 961 NTHDVRTNIVSERRVNSHDFIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
QY 1021 NPKRSNFSSNDHKKNIQYNSRDTKRVREIILSKQNKCNNEYSMEYCTYSDERNSSP 1080
DB 1021 NPKRSNFSSNDHKKNIQYNSRDTKRVREIILSKQNKCNNEYSMEYCTYSDERNSSP 1080
QY 1081 GPCSRERKKLCCQISDYCLKYFNFIYIYNYCIKSEIKSPYKCFKSEQQSSI 1134
DB 1081 GPCSRERKKLCCQISDYCLKYFNFIYIYNYCIKSEIKSPYKCFKSEQQSSM 1134

RESULT 7
Q8MTI6_PLAFA
ID Q8MTI6_PLAFA PRELIMINARY; PRT; 1137 AA.
AC Q8MTI6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte-binding protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5633;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY099889; AAM51625.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 1137 AA; 132348 MW; 30C28B9B376C76A7 CRC64;

Query Match 93.8%; Score 6077; DB 2; Length 1137;
Best Local Similarity 99.6%; Pred. No. 5.1e-247;
Matches 1129; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKGYFNIFLIPLIFLVNIRINESIIIGRTLYNRQDESSDISRVNSPELANNHKTNIYDS 60
DB 1 MKGYFNIFLIPLIFLVNIRINESIIIGRTLYNRQDESSDISRVNSPELANNHKTNIYDS 60

QY 61 DYEDVNNKLINSFVENKSVKKRSLSPINNKTYSYDIIPPSYSYRNDKFNLSNEDNSG 120
DB 61 DYEDVNNKLINSFVENKSVKKRSLSPINNKTYSYDIIPPSYSYRNDKFNLSNEDNSG 120

QY 121 NTSNNFANTSEISIGKDNQYTIQKRTLFCAGIKRKSIKWICRENSKITYCVDPDRK 180
DB 121 NTSNNFANTSEISIGKDNQYTIQKRTLFCAGIKRKSIKWICRENSKITYCVDPDRK 180

QY 181 IQLCVANFLASRLTMEKFEIPLISVNTAKLLYNKNGKDPISFCNELRNSFSDPRS 240
DB 181 IQLCVANFLASRLTMEKFEIPLISVNTAKLLYNKNGKDPISFCNELRNSFSDPRS 240

QY 241 FIGDMDFGGNTDRVKGYINTKFSDDYKKEKVEKLNKIKKEWKEKKNKALNWMHVK 300
DB 241 FIGDMDFGGNTDRVKGYINTKFSDDYKKEKVEKLNKIKKEWKEKKNKALNWMHVK 300

QY 301 NISKECAIIPAEBPQINLWIKWENNFLMEKRLFLNLIKDCVENKKEYACFGGCRLP 360
```

```
DB 301 NISKECAIIPAEBPQINLWIKWENNFLMEKRLFLNLIKDCVENKKEYACFGGCRLP 360
QY 361 SYTSFMKKSKTQMEVLNLYKKKNSGVDKNNFLNDLFKKNKNDLDDFPKNEKEYDDLCD 420
DB 361 SYTSFMKKSKTQMEVLNLYKKKNSGVDKNNFLNDLFKKNKNDLDDFPKNEKEYDDLCD 420
QY 421 CRYTATIISFLNGPAKNDVDIASQINVNDLGRGCGNYKSNNEKSNWCTGTFTNKPFGTC 480
DB 421 CRYTATIISFLNGPAKNDVDIASQINVNDLGRGCGNYKSNNEKSNWCTGTFTNKPFGTC 480
QY 481 EPPRRQTLCLGRYLLHRGHEEDYKEHLIGASYIEAQLLYKYKEKDENALCSIIQNSYA 540
DB 481 EPPRRQTLCLGRYLLHRGHEEDYKEHLIGASYIEAQLLYKYKEKDENALCSIIQNSYA 540
QY 541 DLADIILKSGDIIDKYYGKWEENLNKVKNDKRNESLKI FRKWKWDENKENVWMSAV 600
DB 541 DLADIILKSGDIIDKYYGKWEENLNKVKNDKRNESLKI FRKWKWDENKENVWMSAV 600
QY 601 LKNKETCKDYDFQKIPQFLRMFKWGGDDFCERKEKIYSFESFKVECKKCDENTCKN 660
DB 601 LKNKETCKDYDFQKIPQFLRMFKWGGDDFCERKEKIYSFESFKVECKKCDENTCKN 660
QY 661 KCSEYKWIIDLKSEYEKQVDKYTKDNKKNYNDIDEVKNKEANVYLKEKSKCKOVNFD 720
DB 661 KCSEYKWIIDLKSEYEKQVDKYTKDNKKNYNDIDEVKNKEANVYLKEKSKCKOVNFD 720
QY 721 DKIFNESPNEDMCKCDEIKYLNEIKYPTKHDYIDITFSDTFGDGTPIISINANE 780
DB 721 DKIFNESPNEDMCKCDEIKYLNEIKYPTKHDYIDITFSDTFGDGTPIISINANE 780
QY 781 QQSGKDTNTGNSSETSDPSVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 840
DB 781 QQSGKDTNTGNSSETSDPSVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 840
QY 841 EVHDASTNTGNSSETSDPSVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 900
DB 841 EVHDASTNTGNSSETSDPSVSHPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV 900
QY 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGFDSSRDSSENGRGTTS 960
DB 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGFDSSRDSSENGRGTTS 960
QY 961 NTHDVRTNIVSERRVNSHDFIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
DB 961 NTHDVRTNIVSERRVNSHDFIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020
QY 1021 NPKRSNFSSNDHKKNIQYNSRDTKRVREIILSKQNKCNNEYSMEYCTYSDERNSSP 1080
DB 1021 NPKRSNFSSNDHKKNIQYNSRDTKRVREIILSKQNKCNNEYSMEYCTYSDERNSSP 1080
QY 1081 GPCSRERKKLCCQISDYCLKYFNFIYIYNYCIKSEIKSPYKCFKSEQQSSI 1134
DB 1081 GPCSRERKKLCCQISDYCLKYFNFIYIYNYCIKSEIKSPYKCFKSEQQSSM 1134

RESULT 8
Q8MTI7_PLAFA
ID Q8MTI7_PLAFA PRELIMINARY; PRT; 1137 AA.
AC Q8MTI7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte-binding protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5633;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY099888; AAM51624.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
```

DR GO:0004872; P:receptor activity; IEA.  
DR GO:0009405; P:pathogenesis; IEA.  
DR InterPro: IPR008602; Duffy\_binding.  
DR Pfam: PF05424; Duffy\_binding; 1.  
SQ SEQUENCE 1137 AA; 132320 MW; FEG3C136F697A88A CRC64;

Query Match 93.8%; Score 6077; DB 2; Length 1137;  
Best Local Similarity 99.6%; Pred. No. 5.1e-247;  
Matches 1129; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKGYFNIFLIPLIFLYNIRINESIIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS 60  
Db 1 MKGYFNIFLIPLIFLYNIRINESIIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS 60

Qy 61 DYEDVNNKLNFSFVENSVKKRSLSFINNKTSDYDIIPPSYSYRNDKFNLSNEEDNSG 120  
Db 61 DYEDVNNKLNFSFVENSVKKRSLSFINNKTSDYDIIPPSYSYRNDKFNLSNEEDNSG 120

Qy 121 NTNSNNFANTSEISIGKDNKQYTFIQKRTLFAAGIKRKSIIKWICRENSSEKITVCVPDRK 180  
Db 121 NTNSNNFANTSEISIGKDNKQYTFIQKRTLFAAGIKRKSIIKWICRENSSEKITVCVPDRK 180

Qy 181 IQLCVANPLNSRLTMEKFEIPLISVNTAKLLYNKNEGKPSIFCNELRNSPSPRSS 240  
Db 181 IQLCVANPLNSRLTMEKFEIPLISVNTAKLLYNKNEGKPSIFCNELRNSPSPRSS 240

Qy 241 FIGDDMDPGGNTDRVKGIVNTKPSDYKKNVEKLNIIKKEWKEKYNLANLWNNHMIVNHKG 300  
Db 241 FIGDDMDPGGNTDRVKGIVNTKPSDYKKNVEKLNIIKKEWKEKYNLANLWNNHMIVNHKG 300

Qy 301 NISKECAIIPAEBQIINLWIKENWENFLWEKKRFLNIIKDKCVENKKYEAFCGGRCPSCS 360  
Db 301 NISKECAIIPAEBQIINLWIKENWENFLWEKKRFLNIIKDKCVENKKYEAFCGGRCPSCS 360

Qy 361 SYTSFMKSKTQMEVLTNLYKKKSGVDXNNFLNDFKQNNKNDLDDFFKNEKEYDLDLDC 420  
Db 361 SYTSFMKSKTQMEVLTNLYKKKSGVDXNNFLNDFKQNNKNDLDDFFKNEKEYDLDLDC 420

Qy 421 CRYTATIIKSPFANGPAKNDVDIASQINVDLIRGFCGNYSKNEKSWNCTGFTTNKFPQTC 480  
Db 421 CRYTATIIKSPFANGPAKNDVDIASQINVDLIRGFCGNYSKNEKSWNCTGFTTNKFPQTC 480

Qy 481 EPPRRQTLCLGRTYLLHGHEDYKHEHLGASIIYEAQLLYKKEKDNALCSIIQNSYA 540  
Db 481 EPPRRQTLCLGRTYLLHGHEDYKHEHLGASIIYEAQLLYKKEKDNALCSIIQNSYA 540

Qy 541 DLADIIKGSDDIIKDYGKKMBENLVNKKDKKRNEESLKIIFREKWDENKENVWKMVASV 600  
Db 541 DLADIIKGSDDIIKDYGKKMBENLVNKKDKKRNEESLKIIFREKWDENKENVWKMVASV 600

Qy 601 LKNKETCKDYKPKQIKIPLFWPKWGGDDFCERKEKIIYSPESFKVECKKDCDENTCKN 660  
Db 601 LKNKETCKDYKPKQIKIPLFWPKWGGDDFCERKEKIIYSPESFKVECKKDCDENTCKN 660

Qy 661 KCSYKWKWIDLKSEYKQVDKYTKDKKKMYDNIDEVNKEANVYLKESKECKDVNFD 720  
Db 661 KCSYKWKWIDLKSEYKQVDKYTKDKKKMYDNIDEVNKEANVYLKESKECKDVNFD 720

Qy 721 DKIFNESPNEYDMCKKDEIKYLNEIKYPKTKHDIYDIDTFSDFGDTGPISINANINE 780  
Db 721 DKIFNESPNEYDMCKKDEIKYLNEIKYPKTKHDIYDIDTFSDFGDTGPISINANINE 780

Qy 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSETRGILDINDPSVTNNVN 840  
Db 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSETRGILDINDPSVTNNVN 840

Qy 841 EVHDASTQGSVNTSDITNGHSSSLNRTTNAQDIKIGRSGNSQDNQENSHSSDMSG 900  
Db 841 EVHDASTQGSVNTSDITNGHSSSLNRTTNAQDIKIGRSGNSQDNQENSHSSDMSG 900

Qy 901 SLTTIGQVPSDNTQNTYDSQNPHRDTPNALASLPDDKINEIEGDFSRRDSENGRGTTS 960  
Db 901 SLTTIGQVPSDNTQNTYDSQNPHRDTPNALASLPDDKINEIEGDFSRRDSENGRGTTS 960

Qy 961 NTHDVRRTNIVSRRVNSHDFIRNGMANNNAHQQYITQIENNGIIRGOESAGNSVNYKD 1020  
Db 961 NTHDVRRTNIVSRRVNSHDFIRNGMANNNAHQQYITQIENNGIIRGOESAGNSVNYKD 1020

Qy 1021 NPKRSNFSSENDHKONIQEYNSRDTKRVREIIXLSKQKCNKNNEYSMEYCTYSDERNSSP 1080  
Db 1021 NPKRSNFSSENDHKONIQEYNSRDTKRVREIIXLSKQKCNKNNEYSMEYCTYSDERNSSP 1080

Qy 1081 GPCSRERKKLCCQISDYCLKYFNFYISYIYVNCIKSEIKSPYKCFKSEGOSSI 1134  
Db 1081 GPCSRERKKLCCQISDYCLKYFNFYISYIYVNCIKSEIKSPYKCFKSEGOSSM 1134

RESULT 9  
QSEK8\_PLARE PRELIMINARY; PRT; 1282 AA.  
AC QSEK8;  
DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
DE Erythrocyte invasion ligand BABL/EEA-140 (Fragment).  
OS Plasmodium reichenowi.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5854;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Rayner J.C., Huber C.S., Barnwell J.W.;  
RT "Conservation and divergence in erythrocyte invasion ligands:  
Mol. Biochem. Parasitol. 138:243-247 (2004).  
DR EMBL; AV572433; AAT77188.1; -; Genomic\_DNA.  
DR InterPro: IPR008602; Duffy\_binding.  
DR Pfam; PF05424; Duffy\_binding; 1.  
FT NON TER 1  
SQ SEQUENCE 1282 AA; 148469 MW; 53D601B4AE7CEBCC CRC64;

Query Match 86.5%; Score 5606.5; DB 2; Length 1282;  
Best Local Similarity 82.0%; Pred. No. 3.2e-227;  
Matches 1058; Conservative 50; Mismatches 90; Indels 93; Gaps 2;

Qy 4 YFNIYFLIPLIFLYNIRINESIIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDSYE 63  
Db 1 YFNIYFLIPLIFLYNIRINESIIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDSYE 60

Qy 64 DVNNKLNFSFVENSVKKRSLSFINNKTSDYDIIPPSYSYRNDKFNLSNEEDNSGNTN 123  
Db 61 DVNNKLNFSFVENSVKKRSLSFINNKTSDYDIIPPSYSYRNDKFNLSNEEDNSGNTN 120

Qy 124 SNNPANTSEISIGKDNKQYTFIQKRTLFAAGIKRKSIIKWICRENSSEKITVCVPDRKIQL 183  
Db 121 SNNPANTSEISIGKDNKQYTFIQKRTLFAAGIKRKSIIKWICRENSSEKITVCVPDRKIQL 180

Qy 184 CVANFLNSRLTMEKFEIPLISVNTAKLLYNKNEGKPSIFCNELRNSPSPRSSFIG 243  
Db 181 CVANFLNSRLTMEKFEIPLISVNTAKLLYNKNEGKPSIFCNELRNSPSPRSSFIG 240

Qy 244 DDMDFGNTDRVKGIVNTKPSDYKKNVEKLNIIKKEWKEKYNLANLWNNHMIVNHKNIS 303  
Db 241 DDMDFGNTDRVKGIVNTKPSDYKKNVEKLNIIKKEWKEKYNLANLWNNHMIVNHKNIS 300

Qy 304 KECATIPAEBQIINLWIKENWENFLWEKKRFLNIIKDKCVENKKYEAFCGGRCPSCSYT 363  
Db 301 KECATIPAEBQIINLWIKENWENFLWEKKRFLNIIKDKCVENKKYEAFCGGRCPSCSYT 360

Qy 364 SFMKSKTQMEVLTNLYKKKSGVDXNNFLNDFKQNNKNDLDDFFKNEKEYDLDLDCRY 423  
Db 361 SFMKSKTQMEVLTNLYKKKSGVDXNNFLNDFKQNNKNDLDDFFKNEKEYDLDLDCRY 420

Qy 424 TATIIKSPFANGPAKNDVDIASQINVDLIRGFCGNYSKNEKSWNCTGFTTNKFPQTCVPP 483  
Db 421 TATIIKSPFANGPAKNDVDIASQINVDLIRGFCGNYSKNEKSWNCTGFTTNKFPQTCVPP 480

```
QY 484 RRQTLCLGRVYLLHGHGDEYKEHLGLASIEYQALLKYKKEKDNALCSIIQNSYADLA 543
Db 481 RRQTLCLGRVYLLHGHGHEHYKEHLGLASIEYQALLKYKKEKDNALCSIIQNSYADFA 540
QY 544 DIIKGSIIIDYDYGKWEENLKNVKNKDNKNEESLKIIFREKWDENKENVWKVMSAVLKN 603
Db 541 DIIKGSIDMKDYGKNWEESLKNVKNKDNKNEESLKIIFREKWDENKENVWKVMSAVLKN 600
QY 604 KETCKDYKQKIQPOFLRPFKEWGGDDCEKREKIKYISFESFKVECKKDCDENTCKNKC 663
Db 601 KETCKDYKQKIQPOFLRPFKEWGGDDCEKREKIKYISFESFKVECKKDCDENTCKNKC 660
QY 664 EYKWIIDLKSEYKQVYKTKQKKNKYDNIDEVKNKEANVYLKESKECKDVNFDKI 723
Db 661 EYKWIIDLKSEYKQVYKTKQKKNKYDNIDEVKNKEANVYLKESKECKDVNFDKI 720
QY 724 FNSPNEYEDMCKKCDIKYLNELKYPKTKHDIYDIDTFSDTQDGTPIGINANINEQOS 783
Db 721 FNSPNEYEDMCKKCDIKYLNELKYPKTKHDIYDIDTFSTFGGRTPIGINANINEQON 780
QY 784 GKOTSGTNSGTSVSHSPESDAAINVEKLSGDESSSETRGILDINDPSVTNNV--- 839
Db 781 GKOTSGTNGTET-----ESDAGINVEKLSGDEKSSERREISDIKELSVTSNVNEAS 831
QY 840 ----- 839
Db 832 DALTLETEDGENTLANLRKEDKSLSTKQNTDLNGRSTLEEQVHQTNGSDTGVNVM 891
QY 840 -----NEVHDSANTQGSVNTSDITNGHSESSLNRTTNAQDIKIG 879
Db 892 ESNSSGITGRKKEINIDGHEVHAASNTQSSVNTSDIRKEHSESSLNRTTQDIKIG 951
QY 880 RSGNEQSDNQNSHSDSGSLTIGQVPSDNTQNTYDSQNHPRDTPNALASLPDDKI 939
Db 952 SSGNEQSDNQNSHSDSDSLTIKHPDSEDNTQNTYDLQDSKDTSLTVSSQSDDEV 1011
QY 940 NELEGFSSDSENGRGDTSTNTHDVRTIVSERRVNSHDPIFNGMANNNAHOYITQI 999
Db 1012 NGIGDLSSGDSGGGDTISKTHDVRPMNIVREKHVNNHDFTRSGMNNNAHOYITPN 1071
QY 1000 ENNGIIRQGESAGNSVNYKDNPKRNFSSDNDHKKNIQVNSRDTKRVREIILKSKON 1059
Db 1072 GNNGIIRQGESDGGKYNSEHNRARSNFSKNDHKKNIQVNSRDTKRVREIILKSKON 1131
QY 1060 KQNEYSMVCTYSDERNSSPGPCSRERKKLCCQISDYCLKYPNFYSIYYNCIKSEIK 1119
Db 1132 KQNEYSMVCTYSDERNSSPGPCSRERKKLCCQISDYCLKYPNFYSIYYNCIKSEIK 1191
QY 1120 SPEYKCFKSGOSSIPYFAGGILVILVLLSSASRMKSNEEYDICESNIEATFEENY 1179
Db 1192 SPEYKCFKSGRSNIPYFAGGILVILVLLSSASRLGKSNEEYDICEVNEAFAFEENY 1251
QY 1180 LNKLSRIFNOEVOETNISDYSEYNEKNMY 1210
Db 1252 LNKLSQIFNQEVQETNISDYSEYNEKNMY 1282
RESULT 10
Q8MM04 PLAF 1
ID Q8MM04 PLAF PRELIMINARY; PRT; 616 AA.
AC Q8MM04;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE BAEBL protein (Erythrocyte binding antigen region II) (Fragment).
GN Name=eba-140;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
[1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RL Genetids 163:1327-1336 (2003).
DR EMBL; AF507991; AAM45256.1; -; Genomic DNA.
DR EMBL; AF507993; AAM45258.1; -; Genomic DNA.
DR EMBL; AJ438841; CAD27566.1; -; Genomic DNA.
DR EMBL; AJ438843; CAD27568.1; -; Genomic DNA.
DR EMBL; AF507992; AAM45257.1; -; Genomic DNA.
FT NON_TER 616
SQ SEQUENCE 616 AA; 73429 MW; CC7FDEA3CD5F39F2 CRC64;
Query Match 52.1%; Score 3374; DB 2; Length 616;
Best Local Similarity 100.0%; Pred. No. 6.9e-134;
Matches 616; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 141 QYTFIQKRTLHACGIRKRSIKWICRENSSEKITVCPVDRKIQLCVANFLNSRLTMEKFK 200
Db 1 QYTFIQKRTLHACGIRKRSIKWICRENSSEKITVCPVDRKIQLCVANFLNSRLTMEKFK 60
QY 201 EPLISVNTAEKLLYNNKSGKPSIFCNELRNSFSDFRSFGICDDMDFGGNTDRVKYIN 260
Db 61 EPLISVNTAEKLLYNNKSGKPSIFCNELRNSFSDFRSFGICDDMDFGGNTDRVKYIN 120
QY 261 TFPDYYKKNVKNVKNKNIKEWWEKKNANLNHNMVNHKGNISKECAIIPAEPPQINLWI 320
Db 121 TFPDYYKKNVKNVKNKNIKEWWEKKNANLNHNMVNHKGNISKECAIIPAEPPQINLWI 180
QY 321 KENNENFLMEKRLFLNIKDCKVENKYEACFGGCRLLPCSSYTSFMKKSXTQMEVLTNLY 380
Db 181 KENNENFLMEKRLFLNIKDCKVENKYEACFGGCRLLPCSSYTSFMKKSXTQMEVLTNLY 240
QY 381 KKNKSGVDRKNFLNDLPKKNKNDLDDFPKNEKEYDDLDCCRITATIIKSFNGPAKNDV 440
Db 241 KKNKSGVDRKNFLNDLPKKNKNDLDDFPKNEKEYDDLDCCRITATIIKSFNGPAKNDV 300
QY 441 DIASQNLVNDLRGFGVKNKSNKSNWCTGTFTNKPFGTCPPRROTLCIGRTYLLHRGH 500
Db 301 DIASQNLVNDLRGFGVKNKSNKSNWCTGTFTNKPFGTCPPRROTLCIGRTYLLHRGH 360
QY 501 EEDYKSHLLGASIEYQALLKYKKEKDNALCSIIQNSYADLADIIKGSIIIDYDYGK 560
Db 361 EEDYKSHLLGASIEYQALLKYKKEKDNALCSIIQNSYADLADIIKGSIIIDYDYGK 420
QY 561 EENLNKVNKDKKNEESLKIIFREKWDENKENVWKVMSAVLKNKETCKDYKQKIQPOFL 620
Db 421 EENLNKVNKDKKNEESLKIIFREKWDENKENVWKVMSAVLKNKETCKDYKQKIQPOFL 480
QY 621 RWFKEWGGDDCEKREKIKYISFESFKVECKKDCDENTCKNKCSEYKKNIDLKSEYKQV 680
Db 481 RWFKEWGGDDCEKREKIKYISFESFKVECKKDCDENTCKNKCSEYKKNIDLKSEYKQV 540
QY 681 DKYTKDNKKNKYDNIDEVKNKEANVYLKESKECKDVNFDKIIFNSPNEYEDMCKKCD 740
Db 541 DKYTKDNKKNKYDNIDEVKNKEANVYLKESKECKDVNFDKIIFNSPNEYEDMCKKCD 600
QY 741 IKYLNELKYPKTKHDI 756
Db 601 IKYLNELKYPKTKHDI 616
RESULT 11
Q8MM45 PLAF 1
ID Q8MM45 PLAF PRELIMINARY; PRT; 616 AA.
AC Q8MM45;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE BAEBL protein (Erythrocyte binding antigen region II)
```

```
DE (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Baum J., Thomas A.W., Conway D.J.;
RL MEDLINE=22588500; PubMed=12702678;
RX "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RT Genetics 163:1327-1336(2003).
DR EMBL; AF507981; AA45246.1; -; Genomic DNA.
DR EMBL; AF507983; AA45248.1; -; Genomic DNA.
DR EMBL; AF507984; AA45249.1; -; Genomic DNA.
DR EMBL; AF507985; AA45250.1; -; Genomic DNA.
DR EMBL; AJ438833; CAD27558.1; -; Genomic DNA.
DR EMBL; AJ438834; CAD27559.1; -; Genomic DNA.
DR EMBL; AJ438839; CAD27564.1; -; Genomic DNA.
DR EMBL; AJ438849; CAD27574.1; -; Genomic DNA.
DR EMBL; AF507982; AA45247.1; -; Genomic DNA.
FT NON_TER 1
FT NON_TER 616
SQ SEQUENCE 616 AA; 73456 MW; 6FDE19BPEC982587 CRC64;

Query Match 52.0%; Score 3368; DB 2; Length 616;
Best Local Similarity 99.8%; Pred. No. 1.2e-133;
Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 141 QYTFIQKRTLTHFACGIRKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 200
Db 1 QYTFIQKRTLTHFACGIRKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 60

Qy 201 EIFLISVNTAKLLYNKNEGKPSIFCNELRNSFSDFSGFIDGMDPFGGNTDRVKGIN 260
Db 61 EIFLISVNTAKLLYNKNEGKPSIFCNELRNSFSDFSGFIDGMDPFGGNTDRVKGIN 120

Qy 261 TKFSDDYYKKNVEKLNKIKKEWKEKNKANLWNHMIWNHKNISKECAIIPAEBPQINLWI 320
Db 121 KKFSDYYKKNVEKLNKIKKEWKEKNKANLWNHMIWNHKNISKECAIIPAEBPQINLWI 180

Qy 321 KEWNEFLMEKRLFLNIKDKCVENKYYEACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 380
Db 181 KEWNEFLMEKRLFLNIKDKCVENKYYEACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 240

Qy 381 KKNNSGVDKNNFLNDLPKNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSGFLNGPAKNDV 440
Db 241 KKNNSGVDKNNFLNDLPKNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSGFLNGPAKNDV 300

Qy 441 DIASQINVDLGRFGCNYSKNSKSNKSNCTGTFTNKPFGTCPPRRQTLCLGRTYLLHRGH 500
Db 301 DIASQINVDLGRFGCNYSKNSKSNKSNCTGTFTNKPFGTCPPRRQTLCLGRTYLLHRGH 360

Qy 501 EEDYKEHLLGASIEAQLLYKYKKEKDNALCSIIQNSYADLADIIGSDIIKDYGGKQM 560
Db 361 EEDYKEHLLGASIEAQLLYKYKKEKDNALCSIIQNSYADLADIIGSDIIKDYGGKQM 420

Qy 561 EENLNKYNKDKRNEESLKIIFREKWMDEKENVKMSAVLNKKTCKDYDKPKQIKPOFL 620
Db 421 EENLNKYNKDKRNEESLKIIFREKWMDEKENVKMSAVLNKKTCKDYDKPKQIKPOFL 480

Qy 621 RWFKEWDDPCEKEKEKIEYSPESFKVECKKDCDENTCKNCSYKKKWLIDKSEYEKQV 680
Db 481 RWFKEWDDPCEKEKEKIEYSPESFKVECKKDCDENTCKNCSYKKKWLIDKSEYEKQV 540

Qy 681 DKYTKDKNNKMYDNIDVGNKNEANVYLKESKECKQVNFDDKIPNESPNYEYDCKKCKDE 740
Db 541 DKYTKDKNNKMYDNIDVGNKNEANVYLKESKECKQVNFDDKIPNESPNYEYDCKKCKDE 600
```

```
Qy 741 IKYLNELIKYPTKXHI 756
Db 601 IKYLNELIKYPTKXHI 616

RESULT 12
Q8T9L8 PLAFa PRELIMINARY; PRT; 616 AA.
AC Q8T9L8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Erythrocyte binding protein 2 (BAEBL protein) (Fragment).
GN Name-baeb1;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Oak Knoll;
RX MEDLINE=21674692; PubMed=11814568; DOI=10.1016/S0166-6851(01)00428-5;
RA Narum D.L., Fuhrmann S.R., Lau T., Sim B.K.;
RT "A novel Plasmodium falciparum erythrocyte binding protein-2
(BEP2/BAEBL) involved in erythrocyte receptor binding.";
RL Mol. Biochem. Parasitol. 119:159-168(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY052371; AAL23761.1; -; Genomic DNA.
DR EMBL; AF507979; AA45244.1; -; Genomic DNA.
DR EMBL; AF507980; AA45245.1; -; Genomic DNA.
DR EMBL; AF507978; AA45243.1; -; Genomic DNA.
FT NON_TER 1
FT NON_TER 616
SQ SEQUENCE 616 AA; 419A89BAB4AA1F8F CRC64;

Query Match 52.0%; Score 3367; DB 2; Length 616;
Best Local Similarity 99.7%; Pred. No. 1.4e-133;
Matches 614; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 141 QYTFIQKRTLTHFACGIRKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 200
Db 1 QYTFIQKRTLTHFACGIRKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 60

Qy 201 EIFLISVNTAKLLYNKNEGKPSIFCNELRNSFSDFSGFIDGMDPFGGNTDRVKGIN 260
Db 61 EIFLISVNTAKLLYNKNEGKPSIFCNELRNSFSDFSGFIDGMDPFGGNTDRVKGIN 120

Qy 261 TKFSDDYYKKNVEKLNKIKKEWKEKNKANLWNHMIWNHKNISKECAIIPAEBPQINLWI 320
Db 121 KKFSDYYKKNVEKLNKIKKEWKEKNKANLWNHMIWNHKNISKECAIIPAEBPQINLWI 180

Qy 321 KEWNEFLMEKRLFLNIKDKCVENKYYEACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 380
Db 181 KEWNEFLMEKRLFLNIKDKCVENKYYEACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 240

Qy 381 KKNNSGVDKNNFLNDLPKNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSGFLNGPAKNDV 440
Db 241 KKNNSGVDKNNFLNDLPKNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSGFLNGPAKNDV 300

Qy 441 DIASQINVDLGRFGCNYSKNSKSNKSNCTGTFTNKPFGTCPPRRQTLCLGRTYLLHRGH 500
Db 301 DIASQINVDLGRFGCNYSKNSKSNKSNCTGTFTNKPFGTCPPRRQTLCLGRTYLLHRGH 360

Qy 501 EEDYKEHLLGASIEAQLLYKYKKEKDNALCSIIQNSYADLADIIGSDIIKDYGGKQM 560
Db 361 EEDYKEHLLGASIEAQLLYKYKKEKDNALCSIIQNSYADLADIIGSDIIKDYGGKQM 420

Qy 561 EENLNKYNKDKRNEESLKIIFREKWMDEKENVKMSAVLNKKTCKDYDKPKQIKPOFL 620
Db 421 EENLNKYNKDKRNEESLKIIFREKWMDEKENVKMSAVLNKKTCKDYDKPKQIKPOFL 480
```

```
QY 621 RWPKEWGGDDFCCKRKEKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKKESEYKQV 680
DB 481 RWPKEWGGDDFCCKRKEKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKKESEYKQV 540
QY 681 DKYTKDKNKKQMYDNIDEVKNKEANVYLKESKECKDVNFDKIFNESPNEYEDMCKKCD 740
DB 541 DKYTKDKNKKQMYDNIDEVKNKEANVYLKESKECKDVNFDKIFNESPNEYEDMCKKCD 600
QY 741 IKYLNKIKYPKTKHDI 756
DB 601 IKYLNKIKYPKTKHDI 616

RESULT 13
Q81049_PLAFA
ID Q81049_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q81049;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-140;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ438832; CAD27557.1; -; Genomic_DNA.
DR EMBL; AJ438836; CAD27561.1; -; Genomic_DNA.
DR EMBL; AJ438850; CAD27575.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 616 AA; 73483 MW; F6DE19B8FC9FC344 CRC64;

Query Match 51.9%; Score 3365; DB 2; Length 616;
Best Local Similarity 99.7%; Pred. No. 1.7e-133;
Matches 614; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 141 QYTFIQKRTHLFACGIRKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 200
DB 1 QYTFIQKRTHLFACGIRKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 60
QY 201 EIFLISVNTAKLLYNKNEGKPSIFCNELNSFSDFRNSFSGDMDPGGNTDRVKGYIN 260
DB 61 EIFLISVNTAKLLYNKNEGKPSIFCNELNSFSDFRNSFSGDMDPGGNTDRVKGYIN 120
QY 261 TKFSDYYKKNVEKLNKIKKEWKEKNKANLWNHMIVNHKGNISKECAIIPAEBPQINLWI 320
DB 121 KKFSDYYKKNVEKLNKIKKEWKEKNKANLWNHMIVNHKGNISKECAIIPAEBPQINLWI 180
QY 321 KENNENFLMEKKRLFLNIKDCVENKKEYACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 380
DB 181 KENNENFLMEKKRLFLNIKDCVENKKEYACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 240
QY 381 KKNKSGVDKNNFLNDFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSFNGPAKNDV 440
DB 241 KKNKSGVDKNNFLNDFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSFNGPAKNDV 300
QY 441 DIASQINVDLRGFGCNYSKNNEKSNWCTGFTFNKFPGTCEPPRRQTLCLGRVYLLHRGH 500
DB 301 DIASQINVDLRGFGCNYSKNNEKSNWCTGFTFNKFPGTCEPPRRQTLCLGRVYLLHRGH 360
QY 501 EEDYKEHLHGASIIYEAQLLKYYKKEKDNALCSIIQNSYADLADIIGKSDIIKDYYGKGM 560
DB 361 EEDYKEHLHGASIIYEAQLLKYYKKEKDNALCSIIQNSYADLADIIGKSDIIKDYYGKGM 420
QY 561 EENLKNVKDKKRNEESLKIIFREKWDENKENVWMSAVLNKNETCKDYDKFKQIKPOFL 620
DB 561 EENLKNVKDKKRNEESLKIIFREKWDENKENVWMSAVLNKNETCKDYDKFKQIKPOFL 480
QY 621 RWPKEWGGDDFCCKRKEKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKKESEYKQV 680
```

```
DB 421 EENLKNVKDKKRNEESLKIIFREKWDENKENVWMSAVLNKNETCKDYDKFKQIKPOFL 480
QY 621 RWPKEWGGDDFCCKRKEKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKKESEYKQV 680
DB 481 RWPKEWGGDDFCCKRKEKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKKESEYKQV 540
QY 681 DKYTKDKNKKQMYDNIDEVKNKEANVYLKESKECKDVNFDKIFNESPNEYEDMCKKCD 740
DB 541 DKYTKDKNKKQMYDNIDEVKNKEANVYLKESKECKDVNFDKIFNESPNEYEDMCKKCD 600
QY 741 IKYLNKIKYPKTKHDI 756
DB 601 IKYLNKIKYPKTKHDI 616

RESULT 14
Q8MM74_PLAFA
ID Q8MM74_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q8MM74;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE BAEBL protein (Fragment).
GN Name=eba-140;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF507989; AAM45254.1; -; Genomic_DNA.
DR EMBL; AF507990; AAM45255.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 616 AA; 73525 MW; 2C526A348C4E7771 CRC64;

Query Match 51.9%; Score 3364; DB 2; Length 616;
Best Local Similarity 99.5%; Pred. No. 1.8e-133;
Matches 613; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 141 QYTFIQKRTHLFACGIRKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 200
DB 1 QYTFIQKRTHLFACGIRKRSIKWICRENSKIVCVDPDKIQLCVANFLNSRLTMEKFK 60
QY 201 EIFLISVNTAKLLYNKNEGKPSIFCNELNSFSDFRNSFSGDMDPGGNTDRVKGYIN 260
DB 61 EIFLISVNTAKLLYNKNEGKPSIFCNELNSFSDFRNSFSGDMDPGGNTDRVKGYIN 120
QY 261 TKFSDYYKKNVEKLNKIKKEWKEKNKANLWNHMIVNHKGNISKECAIIPAEBPQINLWI 320
DB 121 RKFSDYYKKNVEKLNKIKKEWKEKNKANLWNHMIVNHKGNISKECAIIPAEBPQINLWI 180
QY 321 KENNENFLMEKKRLFLNIKDCVENKKEYACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 380
DB 181 KENNENFLMEKKRLFLNIKDCVENKKEYACFGGCRLLPCSSYTSFMKSKTQMEVLNLY 240
QY 381 KKNKSGVDKNNFLNDFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSFNGPAKNDV 440
DB 241 KKNKSGVDKNNFLNDFKKNKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSFNGPAKNDV 300
QY 441 DIASQINVDLRGFGCNYSKNNEKSNWCTGFTFNKFPGTCEPPRRQTLCLGRVYLLHRGH 500
DB 301 DIASQINVDLRGFGCNYSKNNEKSNWCTGFTFNKFPGTCEPPRRQTLCLGRVYLLHRGH 360
QY 501 EEDYKEHLHGASIIYEAQLLKYYKKEKDNALCSIIQNSYADLADIIGKSDIIKDYYGKGM 560
DB 361 EEDYKEHLHGASIIYEAQLLKYYKKEKDNALCSIIQNSYADLADIIGKSDIIKDYYGKGM 420
QY 561 EENLKNVKDKKRNEESLKIIFREKWDENKENVWMSAVLNKNETCKDYDKFKQIKPOFL 620
DB 561 EENLKNVKDKKRNEESLKIIFREKWDENKENVWMSAVLNKNETCKDYDKFKQIKPOFL 480
QY 621 RWPKEWGGDDFCCKRKEKIYFSEFVKVECKKDCDENTCKNKCSEYKWKWIDLKKESEYKQV 680
```

Db 481 RWFKEGDDCEKKEKIYSPESFVCKKXDCDENTCKNKSEYKWKWIDLKSEYKQV 540  
 Qy 681 DKYTKDKNKKMYDNIDVKNKEANVYLKESKECKDVNFDKIFNESPNEYEDCKKCKDE 740  
 Db 541 DKYTKDKNKKMYDNIDVKNKEANVYLKESKECKDVNFDKIFNESPNEYEDCKKCKDE 600  
 Qy 741 IKYLNEIKYPTKXKDI 756  
 Db 601 IKYLNEIKYPTKXKDI 616

## RESULT 15

Q8T9N3 PLAPA PRELIMINARY; PRT; 616 AA.  
 AC Q8T9N3: OSMM01;  
 DT 01-JUN-2002 (TremBLrel. 21, Created)  
 DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)  
 DT 01-FEB-2005 (TremBLrel. 29, Last annotation update)  
 DE Erythrocyte binding protein-2 (BABL protein) (Erythrocyte binding  
 antigen region II) (Fragment).  
 GN Name=EBP2; Synonyms=eba-140;  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=3D7;  
 RX MEDLINE=21674692; PubMed=11814568; DOI=10.1016/S0166-6851(01)00428-5;  
 RA Narum D.L., Fuhrmann S.R., Luu T., Sim B.K.;  
 RT "A novel Plasmodium falciparum erythrocyte binding protein-2  
 (EBP2/BABL) involved in erythrocyte receptor binding.";  
 RL Mol. Biochem. Parasitol. 119:159-168(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Mayer D.G., Mu J.-B., Feng X., Su X.-Z., Miller L.H.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]

## NUCLEOTIDE SEQUENCE

Q8T9N3 PLAPA PRELIMINARY; PRT; 616 AA.  
 AC Q8T9N3: OSMM01;  
 DT 01-MAR-2003 (TremBLrel. 23, Created)  
 DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)  
 DE Erythrocyte binding antigen region II (Fragment).  
 GN Name=eba-140;  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC MEDLINE=22588500; PubMed=12702678;  
 RA Baum J., Thomas A.W., Conway D.J.;  
 RT "Evidence for diversifying selection on erythrocyte-binding antigens  
 of Plasmodium falciparum and P. vivax.";  
 RL Genetics 163:1327-1336(2003).  
 DR EMBL: AY030048; AAK55484.1; -; Genomic DNA.  
 DR EMBL: AF507988; AAM45253.1; -; Genomic DNA.  
 DR EMBL: AJ438830; CAD27555.1; -; Genomic DNA.  
 DR EMBL: AJ438831; CAD27556.1; -; Genomic DNA.  
 DR EMBL: AJ438837; CAD27562.1; -; Genomic DNA.  
 DR EMBL: AJ438838; CAD27563.1; -; Genomic DNA.  
 DR EMBL: AJ438842; CAD27567.1; -; Genomic DNA.  
 DR EMBL: AJ438844; CAD27569.1; -; Genomic DNA.  
 DR EMBL: AJ438845; CAD27570.1; -; Genomic DNA.  
 DR EMBL: AJ438846; CAD27571.1; -; Genomic DNA.  
 DR EMBL: AJ438851; CAD27576.1; -; Genomic DNA.  
 DR EMBL: AJ438852; CAD27577.1; -; Genomic DNA.  
 DR EMBL: AJ438853; CAD27578.1; -; Genomic DNA.  
 DR EMBL: AF507986; AAM45251.1; -; Genomic DNA.  
 DR EMBL: AF507987; AAM45252.1; -; Genomic DNA.  
 FT NON TER 1  
 FT NON TER 1  
 FT NON TER 1  
 SQ SEQUENCE 616 AA; 73497 MW; DB9A9BAB4ADF94C CRC64;

Query Match 51.9%; Score 3364; DB 2; Length 616;  
 Best Local Similarity 99.5%; Pred. No. 1.8e-133;  
 Matches 613; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 141 QYTFIQKRTFLFACGKIKRSIKWICRENSKIKTVCVDPDKIQLCVANFLNRLTMEKFK 200  
 Db 1 QYTFIQKRTFLFACGKIKRSIKWICRENSKIKTVCVDPDKIQLCVANFLNRLTMEKFK 60  
 Qy 201 EIFLISVNTAEKLLYNKNEGKDPISIFCNELRNSPFSDFRSSFIGDDMGFGNTDRVKGIN 260

Db 61 EIFLISVNTAEKLLYNKNEGKDPISIFCNELRNSPFSDFRSSFIGDDMGFGNTDRVKGIN 120  
 Qy 261 TKPSDYDYKERNVEKLANNI KKEWMEKNKANLWNHMI VNHKGNISKECAIIPAEBPQINLWI 320  
 Db 121 KKFSDYYKERNVEKLANNI KKEWMEKNKANLWNHMI VNHKGNISKECAIIPAEBPQINLWI 180  
 Qy 321 KENNENFLMEKRLFLNI KDKCVENKKYEACFGGCRLPCCSSYTSFMKSKTKQMEVLNLY 380  
 Db 181 KENNENFLMEKRLFLNI KDKCVENKKYEACFGGCRLPCCSSYTSFMKSKTKQMEVLNLY 240  
 Qy 381 KKNNSGVDKNNFLNDLFKKNNKNDLDDFPKNEKEYDDLDCRYTATIIKSFINGPAKNDV 440  
 Db 241 KKNNSGVDKNNFLNDLFKKNNKNDLDDFPKNEKEYDDLDCRYTATIIKSFINGPAKNDV 300  
 Qy 441 DIASQINVDLRGFCGNYKSNNEKSWNCTGFTNKFPGTCPPRRQTLCLGRTYLLHRGH 500  
 Db 301 DIASQINVDLRGFCGNYKSNNEKSWNCTGFTNKFPGTCPPRRQTLCLGRTYLLHRGH 360  
 Qy 501 EEDYKEHLGASIIYEAOLLYKYKEDENALCSIIQNSYADLADI IKGSDI IKDYVGKQM 560  
 Db 361 EEDYKEHLGASIIYEAOLLYKYKEDENALCSIIQNSYADLADI IKGSDI IKDYVGKQM 420  
 Qy 561 EENLNKYNKDKGRNEESLKI FKEKWMDENKENVKMSAVLNKNETCKDYDKFKI PQFL 620  
 Db 421 EENLNKYNKDKGRNEESLKI FKEKWMDENKENVKMSAVLNKNETCKDYDKFKI PQFL 480  
 Qy 621 RWPKEWGDDCEKKEKIYSPESFVCKKXDCDENTCKNKSEYKWKWIDLKSEYKQV 680  
 Db 481 RWPKEWGDDCEKKEKIYSPESFVCKKXDCDENTCKNKSEYKWKWIDLKSEYKQV 540  
 Qy 681 DKYTKDKNKKMYDNIDVKNKEANVYLKESKECKDVNFDKIFNESPNEYEDCKKCKDE 740  
 Db 541 DKYTKDKNKKMYDNIDVKNKEANVYLKESKECKDVNFDKIFNESPNEYEDCKKCKDE 600  
 Qy 741 IKYLNEIKYPTKXKDI 756  
 Db 601 IKYLNEIKYPTKXKDI 616

## RESULT 16

Q8I787 PLAPA PRELIMINARY; PRT; 616 AA.  
 AC Q8I787;  
 DT 01-MAR-2003 (TremBLrel. 23, Created)  
 DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)  
 DE Erythrocyte binding antigen region II (Fragment).  
 GN Name=eba-140;  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC MEDLINE=22588500; PubMed=12702678;  
 RA Baum J., Thomas A.W., Conway D.J.;  
 RT "Evidence for diversifying selection on erythrocyte-binding antigens  
 of Plasmodium falciparum and P. vivax.";  
 RL Genetics 163:1327-1336(2003).  
 DR EMBL: AJ438840; CAD27565.1; -; Genomic DNA.  
 FT NON TER 1  
 FT NON TER 1  
 FT NON TER 1  
 SQ SEQUENCE 616 AA; 73442 MW; C50CE1C56E605F54 CRC64;

Query Match 51.9%; Score 3363; DB 2; Length 616;  
 Best Local Similarity 99.7%; Pred. No. 2e-133;  
 Matches 614; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 141 QYTFIQKRTFLFACGKIKRSIKWICRENSKIKTVCVDPDKIQLCVANFLNRLTMEKFK 200  
 Db 1 QYTFIQKRTFLFACGKIKRSIKWICRENSKIKTVCVDPDKIQLCVANFLNRLTMEKFK 60  
 Qy 201 EIFLISVNTAEKLLYNKNEGKDPISIFCNELRNSPFSDFRSSFIGDDMGFGNTDRVKGIN 260

```
Db 61 EIFLISVNTAEKLLYNKSGKOPSIFCNELRNSFSDFRNSFI GDDMDFGGNTDRVKGYIN 120
Qy 261 TKPSDYDYKKEKVKELNNIKKEWMEKNKANLWNHMI VNHKGNISKECAIIPAEBPQINLWI 320
Db 121 KTFSDYYKKEKVKELNNIKKEWMEKNKANLWNHMI VNHKGNISKECAIIPAEBPQINLWI 180
Qy 321 KEWNEFLMEKKRLFLNIKDCKVENKKEACFGGCRLP CESSYTSFMKKSXTQMEVLTNLY 380
Db 181 KEWNEFLMEKKRLFLNIKDCKVENKKEACFGGCRLP CESSYTSFMKKSXTQMEVLTNLY 240
Qy 381 KKNKSGVDKKNFLNDLFPKNNKNDLDDFPKNEKE YDLDCCRVTATTIISKFLNGPAKNDV 440
Db 241 KKNKSGVDKKNFLNDLFPKNNKNDLDDFPKNEKE YDLDCCRVTATTIISKFLNGPAKNDV 300
Qy 441 DIASQINVNDLRGFGCNYKSNNEKSNWCTGTFNKF PGTCPPRRQTLCLGRYTYLLHRGH 500
Db 301 DIASQINVNDLRGFGCNYKSNNEKSNWCTGTFNKF PGTCPPRRQTLCLGRYTYLLHRGH 360
Qy 501 EEDYKEHLLGASIEYAQLLKYKYKEKDENALCSII QNSYADLADI IKGSDIIKDYYGKKM 560
Db 361 EEDYKEHLLGASIEYAQLLKYKYKEKDENALCSII QNSYADLADI IKGSDIIKDYYGKKM 420
Qy 561 EENLNKVNKDKRNEESLKI FKEKWDENKENVWMSA VLNKNETCKDYDKFKIQIPQFL 620
Db 421 EENLNKVNKDKRNEESLKI FKEKWDENKENVWMSA VLNKNETCKDYDKFKIQIPQFL 480
Qy 621 RWFKEWGGDFCEKREKIKIYSPESFKVECKKDCD ENTCKNKSEYKKWIDLKSEYKQV 680
Db 481 RWFKEWGGDFCEKREKIKIYSPESFKVECKKDCD ENTCKNKSEYKKWIDLKSEYKQV 540
Qy 681 DKYTKDKNKKMYNDIDEVNKEANVYLKESKECKD VNFDDKIFNESPNYEDMCKKDE 740
Db 541 DKYTKDKNKKMYNDIDEVNKEANVYLKESKECKD VNFDDKIFNESPNYEDMCKKDE 600
Qy 741 IKYLNIEIKYPTKTHDI 756
Db 601 IKYLNIEIKYPTKTHDI 616
```

## RESULT 17

```
Q81788_PLAFA
ID Q81788_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q81788;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-140;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ438835; CAD27560.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 616
SQ SEQUENCE 616 AA; 73526 MW; 311F5009C1744A3F CRC64;
```

Query Match 51.8%; Score 3360; DB 2; Length 616;  
Best Local Similarity 99.4%; Pred. No. 2.7e-133;  
Matches 612; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy 141 QYTFIQKRTHLFACGIRKRSIKWICRENSKIKITCVCPDRKIQLCVANFLNRLTMEKFK 200
Db 1 QYTFIQKRTHLFACGIRKRSIKWICRENSKIKITCVCPDRKIQLCVANFLNRLTMEKFK 60
Qy 201 EIFLISVNTAEKLLYNKSGKOPSIFCNELRNSFSDFRNSFI GDDMDFGGNTDRVKGYIN 260
```

```
Db 61 EIFLISVNTAEKLLYNKSGKOPSIFCNELRNSFSDFRNSFI GDDMDFGGNTDRVKGYIN 120
Qy 261 TKPSDYDYKKEKVKELNNIKKEWMEKNKANLWNHMI VNHKGNISKECAIIPAEBPQINLWI 320
Db 121 KTFSDYYKKEKVKELNNIKKEWMEKNKANLWNHMI VNHKGNISKECAIIPAEBPQINLWI 180
Qy 321 KEWNEFLMEKKRLFLNIKDCKVENKKEACFGGCRLP CESSYTSFMKKSXTQMEVLTNLY 380
Db 181 KEWNEFLMEKKRLFLNIKDCKVENKKEACFGGCRLP CESSYTSFMKKSXTQMEVLTNLY 240
Qy 381 KKNKSGVDKKNFLNDLFPKNNKNDLDDFPKNEKE YDLDCCRVTATTIISKFLNGPAKNDV 440
Db 241 KKNKSGVDKKNFLNDLFPKNNKNDLDDFPKNEKE YDLDCCRVTATTIISKFLNGPAKNDV 300
Qy 441 DIASQINVNDLRGFGCNYKSNNEKSNWCTGTFNKF PGTCPPRRQTLCLGRYTYLLHRGH 500
Db 301 DIASQINVNDLRGFGCNYKSNNEKSNWCTGTFNKF PGTCPPRRQTLCLGRYTYLLHRGH 360
Qy 501 EEDYKEHLLGASIEYAQLLKYKYKEKDENALCSII QNSYADLADI IKGSDIIKDYYGKKM 560
Db 361 EEDYKEHLLGASIEYAQLLKYKYKEKDENALCSII QNSYADLADI IKGSDIIKDYYGKKM 420
Qy 561 EENLNKVNKDKRNEESLKI FKEKWDENKENVWMSA VLNKNETCKDYDKFKIQIPQFL 620
Db 421 EENLNKVNKDKRNEESLKI FKEKWDENKENVWMSA VLNKNETCKDYDKFKIQIPQFL 480
Qy 621 RWFKEWGGDFCEKREKIKIYSPESFKVECKKDCD ENTCKNKSEYKKWIDLKSEYKQV 680
Db 481 RWFKEWGGDFCEKREKIKIYSPESFKVECKKDCD ENTCKNKSEYKKWIDLKSEYKQV 540
Qy 681 DKYTKDKNKKMYNDIDEVNKEANVYLKESKECKD VNFDDKIFNESPNYEDMCKKDE 740
Db 541 DKYTKDKNKKMYNDIDEVNKEANVYLKESKECKD VNFDDKIFNESPNYEDMCKKDE 600
Qy 741 IKYLNIEIKYPTKTHDI 756
Db 601 IKYLNIEIKYPTKTHDI 616
```

## RESULT 18

```
Q81786_PLAFA
ID Q81786_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q81786;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-140;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ438847; CAD27572.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 616
SQ SEQUENCE 616 AA; 73455 MW; 18DE19BFB9C82582 CRC64;
```

Query Match 51.8%; Score 3359; DB 2; Length 616;  
Best Local Similarity 99.7%; Pred. No. 3e-133;  
Matches 614; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy 141 QYTFIQKRTHLFACGIRKRSIKWICRENSKIKITCVCPDRKIQLCVANFLNRLTMEKFK 200
Db 1 QYTFIQKRTHLFACGIRKRSIKWICRENSKIKITCVCPDRKIQLCVANFLNRLTMEKFK 60
Qy 201 EIFLISVNTAEKLLYNKSGKOPSIFCNELRNSFSDFRNSFI GDDMDFGGNTDRVKGYIN 260
```

Db 61 EFLISVNTAEKLLYNKNEGKPSIFCNELRNSPDRSSFGDDMDFGGNTDRVKGYN 120  
Qy 261 TKFSDYYKKNVEKLNINIKKEWKEKNKANLWNHMI VNHKGNISKECAIIPAEEPQINLWI 320  
Db 121 KKFSDYYKKNVEKLNINIKKEWKEKNKANLWNHMI VNHKGNISKECAIIPAEEPQINLWI 180  
Qy 321 KEWNEFLMEKRLFLNLIKORCVENKGYEACFGGCRLPCCSSYTSFMKSKTQMEVLTNLY 380  
Db 181 KEWNEFLMEKRLFLNLIKORCVENKGYEACFGGCRLPCCSSYTSFMKSKTQMEVLTNLY 240  
Qy 381 KKKSGVDKNNFLNDLPKNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSPFLNGPAKNDV 440  
Db 241 KKKSGVDKNNFLNDLPKNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSPFLNGPAKNDV 300  
Qy 441 DIASQINVDLGRFGCNYSNNEKSNWCTGFTNKPFGTCPPRRQTLCLGRTYLLHRGH 500  
Db 301 DIASQINVDLGRFGCNYSNNEKSNWCTGFTNKPFGTCPPRRQTLCLGRTYLLHRGH 360  
Qy 501 EEDYKEHLGLGASIIYEAQLLYKYKEKDENALCSIIQNSYADLADI IKGSDI IKDYGGKQV 560  
Db 361 EEDYKEHLGLGASIIYEAQLLYKYKEKDENALCSIIQNSYADLADI IKGSDI IKDYGGKQV 420  
Qy 561 EENLNKVNKDKKRNESLKIIPREKWNDEKENVKWSAVLNKNETCKDYDKFKQIPQFL 620  
Db 421 EENLNKVNKDKKRNESLKIIPREKWNDEKENVKWSAVLNKNETCKDYDKFKQIPQFL 480  
Qy 621 RWFKEWGDDECEKEKEKLYSPESFKVCKCKKDCDENTCKNCKSKYKWKWIDLKSEYKQV 680  
Db 481 RWFKEWGDDECEKEKEKLYSPESFKVCKCKKDCDENTCKNCKSKYKWKWIDLKSEYKQV 540  
Qy 681 DKYTKDNKKQYDNIDEVKNKEANVYLKESKECKDVNFDDKI FNEPNEVEDMCKKCKDE 740  
Db 541 DKYTKDNKKQYDNIDEVKNKEANVYLKESKECKDVNFDDKI FNEPNEVEDMCKKCKDE 600  
Qy 741 IKYLNIEIKYPTKXHI 756  
Db 601 IKYLNIEIKYPTKXHI 616

## RESULT 19

Q66PM7 PLAF PA PRELIMINARY; PRT; 606 AA.  
AC Q66PM7  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, last annotation update)  
DE Erythrocyte binding protein-2 (Fragment).  
GN Name=EBP-2;  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]

## NUCLEOTIDE SEQUENCE.

RP STRAIN=ALK, BHZ, FFS, GOS, GVM, IPN, JSL, LANA, NSI, RPN, and 21Q;  
RC PubMed=15385490; DOI=10.1128/IAI.72.10.5886-5891.2004;  
RX Lobo C.A., de Frazao K., Rodriguez M., Reid M., Zalis M.,  
RA Lustigman S.;  
RA "Invasion Profiles of Brazilian Field Isolates of Plasmodium  
falciparum: Phenotypic and Genotypic Analyses.";  
RT Infect. Immun. 72:5886-5891 (2004).  
DR EMBL; AY652782; AAT99597.1; -; mRNA.  
DR EMBL; AY652783; AAT99598.1; -; mRNA.  
DR EMBL; AY652784; AAT99599.1; -; mRNA.  
DR EMBL; AY652785; AAT99600.1; -; mRNA.  
DR EMBL; AY652786; AAT99601.1; -; mRNA.  
DR EMBL; AY652787; AAT99602.1; -; mRNA.  
DR EMBL; AY652788; AAT99603.1; -; mRNA.  
DR EMBL; AY652789; AAT99604.1; -; mRNA.  
DR EMBL; AY652790; AAT99605.1; -; mRNA.  
DR EMBL; AY652793; AAT99608.1; -; mRNA.  
DR EMBL; AY652780; AAT99595.1; -; mRNA.  
FT NON\_TER

FT NON\_TER 606 606  
SQ SEQUENCE 606 AA; 72126 MW; 91B73964716D7292 CRC64;  
Query Match 51.2%; Score 3319; DB 2; Length 606;  
Best Local Similarity 100.0%; Pred. NO. 1.4e-131;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 151 LFACGIRKRSIKWICRENSBKITVCVDPDKIQLCVANFLNSRLTMEKFEIFLISVNT 210  
Db 1 LFACGIRKRSIKWICRENSBKITVCVDPDKIQLCVANFLNSRLTMEKFEIFLISVNT 60  
Qy 211 AKLLYNKNEGKPSIFCNELRNSPDRSSFGDDMDFGGNTDRVKGYNITKFSDYK 270  
Db 61 AKLLYNKNEGKPSIFCNELRNSPDRSSFGDDMDFGGNTDRVKGYNITKFSDYK 120  
Qy 271 NVEKLNNIKKEWKEKNKANLWNHMI VNHKGNISKECAIIPAEEPQINLWIKENENFLME 330  
Db 121 NVEKLNNIKKEWKEKNKANLWNHMI VNHKGNISKECAIIPAEEPQINLWIKENENFLME 180  
Qy 331 KGRFLNLIKDCVENKGYEACFGGCRLPCCSSYTSFMKSKTQMEVLTNLYKKNKSGVDKN 390  
Db 181 KGRFLNLIKDCVENKGYEACFGGCRLPCCSSYTSFMKSKTQMEVLTNLYKKNKSGVDKN 240  
Qy 391 NFNLDLPKNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSPFLNGPAKNDVDIASQINVD 450  
Db 241 NFNLDLPKNNKNDLDDFFKNEKEYDDLCDCRYTATIIKSPFLNGPAKNDVDIASQINVD 300  
Qy 451 LRFGGCHYKNNKSNWCTGFTNKPFGTCPPRRQTLCLGRTYLLHRGHEDYKEHLG 510  
Db 301 LRFGGCHYKNNKSNWCTGFTNKPFGTCPPRRQTLCLGRTYLLHRGHEDYKEHLG 360  
Qy 511 ASIYEAQLLYKYKEKDENALCSIIQNSYADLADI IKGSDI IKDYGGKWEENLNKVKND 570  
Db 361 ASIYEAQLLYKYKEKDENALCSIIQNSYADLADI IKGSDI IKDYGGKWEENLNKVKND 420  
Qy 571 KGRNEESLKIIPREKWNDEKENVKWSAVLNKNETCKDYDKFKQIPQLRWFKEWGD 630  
Db 421 KGRNEESLKIIPREKWNDEKENVKWSAVLNKNETCKDYDKFKQIPQLRWFKEWGD 480  
Qy 631 CEKREKEKLYSPESFKVCKCKKDCDENTCKNCKSKYKWKWIDLKSEYKQVDTKDKKK 690  
Db 481 CEKREKEKLYSPESFKVCKCKKDCDENTCKNCKSKYKWKWIDLKSEYKQVDTKDKKK 540  
Qy 691 MYDNIDEVKNKEANVYLKESKECKDVNFDDKI FNEPNEVEDMCKKCKDEIKYLNIEIKY 750  
Db 541 MYDNIDEVKNKEANVYLKESKECKDVNFDDKI FNEPNEVEDMCKKCKDEIKYLNIEIKY 600  
Qy 751 KTKXHI 756  
Db 601 KTKXHI 606

## RESULT 20

Q66PL8 PLAF PA PRELIMINARY; PRT; 606 AA.  
AC Q66PL8  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, last annotation update)  
DE Erythrocyte binding protein-2 (Fragment).  
GN Name=EBP-2;  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]

## NUCLEOTIDE SEQUENCE.

RP STRAIN=04Q, PSS1, and 350;  
RC PubMed=15385490; DOI=10.1128/IAI.72.10.5886-5891.2004;  
RX Lobo C.A., de Frazao K., Rodriguez M., Reid M., Zalis M.,  
RA Lustigman S.;  
RA "Invasion Profiles of Brazilian Field Isolates of Plasmodium  
falciparum: Phenotypic and Genotypic Analyses.";  
RT Infect. Immun. 72:5886-5891 (2004).  
FT NON\_TER

```
DR EMBL; AY652791; AAT99606.1; -; mRNA.
DR EMBL; AY652792; AAT99607.1; -; mRNA.
DR EMBL; AY652791; AAT99596.1; -; mRNA.
FT NON_TER 1
FT NON_TER 606
SQ SEQUENCE 606 AA; 72194 MW; 85526E7D089FB22C CRC64;

Query Match 51.1%; Score 3309; DB 2; Length 606;
Best Local Similarity 99.5%; Pred. No. 3.6e-131;
Matches 603; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 151 LFACGIRKRIKIKWCNSENKIVCVDPDRKIQLCVANFLNSRLTMEKFEIFLISVNT 210
DB 1 LFACGIRKRIKIKWCNSENKIVCVDPDRKIQLCVANFLNSRLTMEKFEIFLISVNT 60

QY 211 AKLLYNKEGKDPISFCNELRNSFSDFRSSFIGDDMDFGGNTDRVKGYINTKFSDDYKEK 270
DB 61 AKLLYNKEGKDPISFCNELRNSFSDFRNSFIGDDMDFGGNTDRVKGYINTKFSDDYKEK 120

QY 271 NVEKLNNIKKEWWEKKNLANLWNHMIWNHKNISKECAIIPAEBPQINLWIKENWENFLME 330
DB 121 NVEKLNNIKKEWWEKKNLANLWNHMIWNHKNISKECAIIPAEBPQINLWIKENWENFLME 180

QY 331 KKRLLFLNIDKCVENKYYEACFGGCRLLPCSSYTSFMKSKTOMEVLTNLYKKNSGVYDKN 390
DB 181 KKRLLFLNIDKCVENKYYEACFGGCRLLPCSSYTSFMKSKTOMEVLTNLYKKNSGVYDKN 240

QY 391 NFLNDLFKKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSLFNGPAKNDVDIASQINVD 450
DB 241 NFLNDLFKKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSLFNGPAKNDVDIASQINVD 300

QY 451 LRFGGCHYKNNKSNWCTGTFTNKPFGTCPPRRQTLCLGRVYLLHRGHEEDYKEHLG 510
DB 301 LRFGGCHYKNNKSNWCTGTFTNKPFGTCPPRRQTLCLGRVYLLHRGHEEDYKEHLG 360

QY 511 ASIYEAOQLLYKYEKEDENALCSIIQNSYADLADIIGSDIIIDKYCGKMEENLNKYNKD 570
DB 361 ASIYEAOQLLYKYEKEDENALCSIIQNSYADLADIIGSDIIIDKYCGKMEENLNKYNKD 420

QY 571 KGRNEESLKIIFREKWDENKENVWKMVASVLKNKETCKDYDKFQKIPOFLRWFKWGDDF 630
DB 421 KGRNEESLKIIFREKWDENKENVWKMVASVLKNKETCKDYDKFQKIPOFLRWFKWGDDF 480

QY 631 CERKEKIIYFESFKVECKKDCDENTCKNKCSEYKWKIIDLKSEYKQVDYTKDQKK 690
DB 481 CERKEKIIYFESFKVECKKDCDENTCKNKCSEYKWKIIDLKSEYKQVDYTKDQKK 540

QY 691 MYDNIDEVKNKEANVYLKESKECKVDNFDDKIFNESPNEDMCKKDEIKYLNEIKYP 750
DB 541 MYDNIDEVKNKEANVYLKESKECKVDNFDDKIFNESPNEDMCKKDEIKYLNEIKYP 600

QY 751 KTKHDI 756
DB 601 KTKHDI 606

RESULT 21
Q81760 PLARE PRELIMINARY; PRT; 616 AA.
AC Q81760;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Putative erythrocyte binding antigen region II (fragment).
GN Name-eba-140 homologue;
OS Plasmodium reichenowi;
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens

RT of Plasmodium falciparum and P. vivax.";
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ438829; CAD27554.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 616
SQ SEQUENCE 616 AA; 73518 MW; DAA91E321E6110E1 CRC64;

Query Match 48.5%; Score 3144; DB 2; Length 616;
Best Local Similarity 92.2%; Pred. No. 3.1e-124;
Matches 568; Conservative 20; Mismatches 28; Indels 0; Gaps 0;

QY 141 QYTFIOKRTHLFACGIRKRSIKWICRENSKIVCVDPDRKIQLCVANFLNSRLTMEKFK 200
DB 1 QYTFIOKRTHLFACGIRKRSIKWICRENSKIVCVDPDRKIQLCVANFLNSRLTMEKFK 60

QY 201 EIFLISVNTAEKLLYNKNEGKDPISFCNELRNSFSDFRSSFIGDDMDFGGNTDRVKGYIN 260
DB 61 EIFLISVNTAEKLLYNKNEGKDPISFCNELRNSFSDFRSSFIGDDMDFGGNTDRVKGYIN 120

QY 261 TKFSDYYKKNVEKLANNIKEMWEKKNLANLWNHKNISKECAIIPAEBPQINLWI 320
DB 121 MKFSDYYAEKQVEKLANNIKEMWEKKNLANLWNHKNISKECATIPEESEPQINAWI 180

QY 321 KENNENFLMEKRLFLNIDKCVENKYYEACFGGCRLLPCSSYTSFMKSKTOMEVLTNLY 380
DB 181 KENNENFLMEKRLFLNIDKCVENKYYEACFGGCRLLPCSSYTSFMKSKTOMEVLTNLY 240

QY 381 KKNKSGVDKNNFLNDLFPKKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSLFNGPAKNDV 440
DB 241 KKNKSGVDKNNFLNDLFPKKNKNDLDDFFKNEKEYDDLCDCRYTATIIKSLFNGPAKNDV 300

QY 441 DIASQINVDLRFGGCHYKNNKSNWCTGTFTNKPFGTCPPRRQTLCLGRVYLLHRGH 500
DB 301 DTASKINVDLRVFGGCHYKNNKSNWCTGTFTNKPFGTCPPRRQTLCLGRVYLLHRGH 360

QY 501 EEDYKEHLGASIIYEAOQLLYKYEKEDENALCSIIQNSYADLADIIGSDIIIDKYCGK 560
DB 361 EEHYKEHLGASIIYEAOQLLYKYEKEDENALCSIIQNSYADLADIIGSDIIIDKYCGK 420

QY 561 EENLKNVKNDKKNNEESLKIIFREKWDENKENVWKMVASVLKNKETCKDYDKFQKIPOFL 620
DB 421 EESLKNVKNDKKNNEESLKIIFREKWDENKENVWKMVASVLKNKETCKDYDKFQKIPOFL 480

QY 621 RWPKEGDDPFCEKREKIIYFESFKVECKKDCDENTCKNKCSEYKWKIIDLKSEYKQV 680
DB 481 RWPKEGDDPFCEKREKIIYFESFKVECKKDCDENTCKNKCSEYKWKIIDLKSEYKQV 540

QY 681 DKYTKDKNNKKNVNDIDEVKNKEANVYLKESKECKVDNFDDKIFNESPNEDMCKKDE 740
DB 541 EKYTKDKNNNIYDNIDEVKNKEANVYLKESKECKVDNFDDKIFNESPNEDMCKKDE 600

QY 741 IKYLNKIKYPTKXDI 756
DB 601 IKYLNKIKYPTKXDI 616

RESULT 22
Q25842 PLAPA PRELIMINARY; PRT; 1475 AA.
AC Q25842;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Erythrocyte binding protein.
GN Name-erythrocyte-binding protein;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92357776; PubMed=1496004;
RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
RT "A family of erythrocyte binding proteins of malaria parasites."
```





FT REGION 159 1104 Essential for binding to erythrocytes.  
 FT VARIANT 1031 1031 E -> V (in strain PCR-3 and strain ITG).  
 SQ SEQUENCE : 1435 AA; 167390 MW; 3244309021B1C3D6 CRC64;

Query Match 24.0%; Score 1557; DB 1; Length 1435;  
 Best Local Similarity 28.5%; Pred. No. 3.2e-57;  
 Matches 410; Conservative 214; Mismatches 467; Indels 348; Gaps 42;

QY 1 MKGYFNIYFLIPLIFLYNIRINESIIGRTLYNRQDESSDIISRVNSPELNHNHKTNYDS 60  
 DB 1 MKNISIFYPFVLYFARNEYDI-----KENEFLDVYKEKFNELDKKYGKQVKT 55

QY 61 DYEDVNNKLINSFVENKSVKKRSLSPINN-----KTKSY---DIIPPSYRNKDNKNSL 112  
 DB 56 D-----KKIIFTIENK-----LDILNNSKFNKRKWSYGTDPNIDKNMSLIN----- 96

QY 113 SENEDSGNTSNFANTSEISIGKQKQ-----YTFIQKTH----- 150  
 DB 97 KHNEEMFNYYQSLSTS--SLIKQNKYVPINAVRVSRILSFLDSRINNGRNTSSNEV 154

QY 151 LFACGIGKRKSIKWCIRENSEKIT--VCVPDRKIQLCVANFLNRLSETEKPKKEIFLISVNT 209  
 DB 155 LSNCREKRGKWKCKKNDNSVVCIPDRIQICIVNLISIIKTYTETKMDHIFIAKX 214

QY 210 EAKLLYNKNEGKPSIFCNELRNFSDFRSFIDGDMDFGGTDRVKGYINTKFSDYKX 269  
 DB 215 ESQALLKNDKNKNSKFCNDLKNLFLDYGLHMGNDMDFGGYSTKAENKIQEVFKGAHGE 274

QY 270 KAVEKLANIKKEWEKAKANLWNIWVNHKNISKECAIPAEPSQNLMTKEWNEFLM 329  
 DB 275 ISEHKINFRKWBWNEFREKLWEAMLSHKNNIN--NCKNIPQESLOITOWIKWHGEBLL 333

QY 330 EKKRLFLNIKDKVENKYEACGGCRLLPCSSYTSFMKCKSKTOMEVLNTLYKKNKSGVDK 389  
 DB 334 ERNRSKLPKSKKNNTLYEACEKECDPCMKYRDWIIIRKFPWHTUSKEYETOK--VPK 391

QY 390 NNFLNDLFLK--KNRQND--LDQFFKN--EKESYDLDCCRYTATIIKSLFLNGPAKNDVDIASQ 445  
 DB 392 ENAENYLIKISEKNDKAVSILLNCAEYSKYCDCKHTTLVKSVLNGNDNTIKEREH 451

QY 446 INVNDLRGFGNYKS--NNEKSNWCTGTFNNKFPCTCEPRPQPLCLGRVYLLHRGHEED 503  
 DB 452 IDLDQDFKFGCDKNSVDNTKVMWECQNPYILSTKDVCPVPRRQBELCLGNIDRIYDKNLLM 511

QY 504 YKEHLGASIVYEAOLLYKYEKEDENALCSTIQNSYADLADIIGKSDIILDYKCKKEEN 563  
 DB 512 IKEHILAIVESILKRGKAKKQDDKEVKLINKTFADIRDIIGGTQWMDLSNRKLGVK 571

QY 564 LNKVNDKGRNBSLKIIFREKWDENKENYKVMASVAVLNKNETCKDYDKFKQIPQLRWF 623  
 DB 572 INTNSKYVHRKNDKLFREEMKVIKDVVNVISWVFKDVTCKE--DDIENIPQPFRRF 630

QY 624 KEWGDPCERKKEKIIYFSPKVCCKKDDCENCKKCKSEYKQWIDLKSEYKQV--- 680  
 DB 631 SEWGGDYCODTKMI-----ETLKVCEKPCEDDDNCKSKNSYKWKISKKKEEYKQAKQY 687

QY 681 DKYTKDKKMYONIDEVKKEANVLYKESKECKOVNEDDKIWNESPNVEDCKKCDR 740  
 DB 688 QEYQKGNYYKMY-----SEFKSIKPEVLYLKYSKCSNLNFEDEFKELHSDYKKNCTWCP 744

QY 741 IKYL----- 746  
 DB 745 VKDVPISIRNEQTSQEAPEPENTEIAHRTETPSISEGPKGNQKERDDDSLKISVSP 804

QY 747 -----IKYPT----- 752  
 DB 805 ENSRPETDAKOTSNLLKLGVDVDSMPKAVIGSSPDNINVTQGDNISGVNSKPLSDDV 864

QY 753 ---KHDIYDID-----TFSTFDGTPIST----- 774  
 DB 865 RPKCKELEDQNSDESETVVNHISKSPSINNDDSGSGSATVSSSSNTGLSDDDRNG 924

QY 775 -----NANINEQSGKOTSGTNSGTSDSP---VSHEPESDAAINV 812

DB 925 DTFVTRQDTANTEDVIRKENADKDEKGADEERHSTSESLSPEERKMLTNEGNSLNH 984  
 QY 813 EKL-----SGDESSETRGILDIN-----DPSVTNNVNEVH-----DAS 846  
 DB 985 EEVKEHTNSNDNVQSGIYVNMVNEKELKOTLENPSSSLDEGKAHEELSEPNLSSDQDMS 1044

QY 847 NTQGSVNTSITDN---GHSESSLNR-----TTNAQDIKIGRSGNEQSDN---OENS 892  
 DB 1045 NTGCLDNTSEETTERISNNEYKVNREDERTLKEYEDIVLAKSHMRRESDDGELYDENS 1104

QY 893 SHSDNSGSLTI--GVPSSEDNTQNTY-----DSQNPHRDTNALA-----SLP- 934  
 DB 1105 DLSTVNDSEDAKAKMGNDTSEMHSNSSHIESDQCKDMKVTGDLGTTTHVONEISVPV 1164

QY 935 ---SDDKINE-----LEGFSSRDSRSENGRDTTSTNTHDYR-----R 967  
 DB 1165 TGEIDEXLRESKESKIHKAEEERLSHTDIHKINPEDRNSNTLHLKDIRNEENERHLTNQN 1224

QY 968 TNTVSERRVNSHDFIRNGMANNNAHOYITQ-----IENNGIIRGOESAGNSVNYKDN- 1021  
 DB 1225 INTSQERDLQKHGF-----HTMNLHGGVRSERQINSHHGNRQDRGNSGVNLNMSNN 1280

QY 1022 -----PKRSNFSENDDKXIOEYNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDE 1075  
 DB 1281 NFNFNIPSRNVL--YDKLDDLLENRNDSTTKELIKKLAEINKCENEISVKYCDMH 1337

QY 1076 RNSSPGSPCSREERKLCQISDYCLKYFNFSYBYNCIKSEIKSPYKCFKSGQSSI 1134  
 DB 1338 EEIPLKTCCTKEKTRNLCCAVSDYCMYFTYDSEYNTCTKREDFDPSYTCFRKEAFSSM 1396

RESULT 25  
 Q9NG63\_PLAFA PRELIMINARY; PRT; 1435 AA.  
 AC Q9NG63;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Erythrocyte binding antigen 175.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RA NUCLEOTIDE SEQUENCE.  
 RA Ma C.L., Yu X.B., Li X.R., Shan Z.X.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF258781; AAF72186.1; -; Genomic DNA.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0009405; P:pathogenesis; IEA.  
 DR InterPro; IPR008602; Duffy\_binding.  
 DR Pfam; PF05424; Duffy\_binding; 2.  
 SQ SEQUENCE 1435 AA; 167269 MW; PF958828C7794F22 CRC64;

Query Match 23.9%; Score 1547; DB 2; Length 1435;  
 Best Local Similarity 28.4%; Pred. No. 8.4e-57;  
 Matches 410; Conservative 216; Mismatches 460; Indels 358; Gaps 44;

QY 1 MKGYFNIYFLIPLIFLYNIRINESIIGRTLYNRQDESSDIISRVNSPELNHNHKTNYDS 60  
 DB 1 MKNISIFYPFVLYFARNEYDI-----KENEFLDVYKEKFNELDKKYGKQVKT 55

QY 61 DYEDVNNKLINSFVENKSVKKRSLSPINN-----KTKSY---DIIPPSYRNKDNKNSL 112  
 DB 56 D-----KKIIFTIENK-----LDILNNSKFNKRKWSYGTDPNIDKNMSLIN----- 96

QY 113 SENEDSGNTSNFANTSEISIGKQKQ-----YTFIQKTH----- 150  
 DB 97 KHNEEMFNYYQSLSTS--SLIKQNKYVPINAVRVSRILSFLDSRINNGRNTSSNEV 154

QY 151 LFACGIGKRKSIKWCIRENSEKIT--VCVPDRKIQLCVANFLNRLSETEKPKKEIFLISVNT 209

Db 155 LSNCREKRGKMGKWDCKKKNDRSNVVCIIPDRRIQICIVNLISIIKTYTETWTDKDFHFIASKK 214  
Qy 210 EAKLLYNKNEGKDPISFCNELRNSFSDFRSSFIGDDMDFGNTDRVKGYNITKFSDDYIKE 269  
Db 215 ESQLLKKNDKNYNSKFCNDLKNISFLDYGHLAGNDMDFGYSTKAENKIQEVFKGAHGK 274  
Qy 270 KNYEKLANNIKKEWEKKNANLWMMIIVNHKNISKECAIIPAEBPQINLWIKENENFLM 329  
Db 275 ISEHILNFKRKNWNEFREKLWEAMLSEHKNNIN-NCKNIPQEBELQITQWIKWHGEBLL 333  
Qy 330 EKKRLFLNIDKQVENKKEACFCGCRLLPCSSYTSFMKSKTOMELVNLTKKNSGVDK 389  
Db 334 ERYNRSKLPKSKCKNNTLYEACEKICDPCWKYRDMIIIRSKFEWHTLSKEYETQN--VSK 391  
Qy 390 NNFLNDLFLK-KNNKND--LDDFFKN-EKEYDDLCDCRYTATIIKSFNLGPAKNDVDIASQ 445  
Db 392 ENAEVYLKISKNKNDKAVSILLNNDCAEYSKYCDCKHHTTLVKSVLNGNDNTIKEKREH 451  
Qy 446 INVNDLRGPGCNYS--NNEKSNWCTGTFNKPFGTCPPRRQTLCLGRYLLHRGHEED 503  
Db 452 IDLDDFKGFCGDKNSVDNTNTKWECKPYILSTKDVCPVPRRQBLCLGNIDRIYDKNLLM 511  
Qy 504 KYEHLGASTYEAOLLYKVKEDENALCSTIQNSYADLADIIGKSDI IKDYCYKKWEN 563  
Db 512 IKEHILAIAYESILKRRKYNKDDKEVKLIINKTFADIRDIIGDTYWNLDLSNRKLVGK 571  
Qy 564 LNKVNKKRNEESLKIIFREKWNDEKENYKVMASVAVLNKNETKOYDKFKQPKQFLRW 623  
Db 572 INTNSVYVHRKNDKLFREWKVVKVVDVWVSWFKDVTCKE--DDIENIPQPPRWF 630  
Qy 624 KENGDDCEKKEKIKYFESFKVECKKDCDENTCKKCSYKKNWIDLKSEYKQV--- 680  
Db 631 SEWDDYCDQKTKMI---ETLKECKEPCEDDDNCKNSKYKESWISKEEYKQAKQY 687  
Qy 681 DKYTKDNKKMYDNIIDVONKEANVYLKESKECKDVNDDKIPNESPNYEDWCKKCD 740  
Db 688 QEYQKGNYYMY---SEPKSIKPVYLYKYSEKSNLNFDEPFKEELHSDYKKNCTWCPE 744  
Qy 741 IKYL-----NS----- 746  
Db 745 VKDVPISIRNEQTSQEAPEENTEITHRTPTPISSEKPGNEQKERRDDSLSKISVSP 804  
Qy 747 -----IKPKT----- 752  
Db 805 ENSRPETDAKOTSNLKLKGDVDISMPKAVIGSSPNNDINVTQGDNISGVNSKPLSDDV 864  
Qy 753 ---KHDYDID-----TFSDTFDGTPTSI----- 774  
Db 865 RPKKELEDQNSDESEETVNVNHSKSPSINNGDSSGSGSATVSSSSNTGLSIDDDRNG 924  
Qy 775 -----NANINEQSGKDTNNTGNETSDSP---VSHEPESDAAINV 812  
Db 925 DTFVRTQDTANTEDVIRKENADKDEKGADEERHSTSESLSSPEEKWSTDEGGSNLN 984  
Qy 813 EKL-----SGDESSETGILDIN-----DPSVTNANVEH-----DAS 846  
Db 985 EEVKEHASNSDNVQSGGIVNMVYKELKOTLENPSSSLDEGKAHEELSEPNLSSDQDMS 1044  
Qy 847 NTQGSVNTSDITN---CHSESSLNR-----TTNAQDIIKGRSGNEQSDN---QENS 892  
Db 1045 NTPGLNTSEETTERISNNEYKVNREGERTLTKEYEDIVLKSHMRESDDGELYDENS 1104  
Qy 893 SHSSDNGSLTI-QQVPSDNTQNTY-----DSQNPHRDTPNALA-----SLP- 934  
Db 1105 DLSTVNDSEDAEAKMGNDTSEMHSNQSIESDQKNDMKTVDGLTTHVQNEISVPV 1164  
Qy 935 ---SDDKINE-----LEGFDSRDSSENGDGTTSNTHDVR-----R 967  
Db 1165 TGEIDELKRESKESKIHABEERLSHTDIIHKINPEDRNSNTLHLKOTIRNEENERHLLTNQ 1224  
Qy 968 TNIVSERRVNSHDF--JRN-----GWAN-----NNAHQVYITQIENNNGIIRGOESAGNSVN 1017  
Db 1225 INTSQERDLQKHGFHTVNNLHGDGVRSRQINSHH-----GNRQRGSGNGVNL 1275

Qy 1018 YKDN-----PKRSNFSSENDHKQIQEYNSRDTKRVREBIIKLSQKQKNNEYSMEYC 1070  
Db 1276 MRSNNNNFNPIRNYL---YDKKLDSDLYENRNDSTTKRELKLAELKNEISEISVKYC 1332  
Qy 1071 TYSDESNSPGCPSPRERKKLCCQISDYCLKYNFYSIYYNCI KSEIKSPYKCFKSEG 1130  
Db 1333 DHMIEEIPLKTKTEKTRNLCAVSDYCMYSFTYDSEYYNCTKREDFDPSYTCRKEA 1392  
Qy 1131 QSSI 1134  
Db 1393 FSSM 1396  
RESULT 26  
Q05644\_PLAPA PRELIMINARY; PRT; 1421 AA.  
AC Q05644;  
DT 01-NOV-1996 (TEMBLrel. 01, Created)  
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
DE Erythrocyte-binding antigen-175 (Fragment).  
GN Name=EBA-175;  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=3D7;  
RA Daugherty J.R., Lanar D.E.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U32207; AAA75179.1; -; Genomic DNA.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR008602; Duffy-binding.  
DR Pfam; PF05424; Duffy-binding; 2.  
FT NON TER 1421 1421  
SQ SEQUENCE 1421 AA; 165330 MW; 7BE8DDFA07CBE771 CRC64;

Query Match 23.7%; Score 1537.5; DB 2; Length 1421;  
Best Local Similarity 28.1%; Pred.No. 2.1e-56;  
Matches 411; Conservative 211; Mismatches 467; Indels 375; Gaps 44;  
Qy 1 MKGYFIYLIPLIFLYNVIRINESIIGRTLYNRQDESSDSRVNSPELNNHKTNIYDS 60  
Db 1 MKCNSIYFPASFPVLYFAKARNEYDI-----KENEKFLDYKKEKNELDKKYGVNQKT 55  
Qy 61 DYEDVANKLINSVENKSVKKGKSLSFINN-----KTKSY---DIIPPSYSYRNDKFNSL 112  
Db 56 D-----KKIFTPIENK-----LDILNNSKFNKRWKSYGTPDNIDKNMSLIN----- 96  
Qy 113 SENEDSGNTNSNPNANTSEISIGDKNQ-----YTFIQKETH----- 150  
Db 97 KHNEEMFNMYOSFLSTS--SLIKQNKYVPINAVRSRILSFLDSRINNGRNTSSNEV 154  
Qy 151 LFAAGIKRKSIIKICHENSEKIT-VCVPDRKIQLCVANFLNSLETWKEFKELISVNT 209  
Db 155 LSNCREKRGKMGKWDCKKKNDRSNVVCIIPDRRIQICIVNLISIIKTYTETWTDKDFHFIASKK 214  
Qy 210 EAKLLYNKNEGKDPISFCNELRNSFSDFRSSFIGDDMDFGNTDRVKGYNITKFSDDYIKE 269  
Db 215 ESQLLKKNDKNYNSKFCNDLKNISFLDYGHLAGNDMDFGYSTKAENKIQEVFKGAHGK 274  
Qy 270 KNYEKLANNIKKEWEKKNANLWMMIIVNHKNISKECAIIPAEBPQINLWIKENENFLM 329  
Db 275 ISEHILNFKRKNWNEFREKLWEAMLSEHKNNIN-NCKNIPQEBELQITQWIKWHGEBLL 333  
Qy 330 EKKRLFLNIDKQVENKKEACFCGCRLLPCSSYTSFMKSKTOMELVNLTKKNSGVDK 389  
Db 334 ERYNRSKLPKSKCKNNTLYEACEKICDPCWKYRDMIIIRSKFEWHTLSKEYETQN--VSK 391  
Qy 390 NNFLNDLFLK-KNNKND--LDDFFKN-EKEYDDLCDCRYTATIIKSFNLGPAKNDVDIASQ 445

```
Db 392 ENAENYLKISGNKNDKAVSLLNCCDAEYSKYCDCKHTTTLVKSVLNGNDNTIKKREH 451
Qy 446 INVNDLRFGGCMYK--NNEKSWNCTGFTWKPFGTCPPRPTCLGRTYLLHRGHEED 503
Db 452 IDLDDFSFGGCKNSVDNTNTWVECKKPKYKLSKTDVCPVRORBLCLGNIDRIYDKNLLM 511
Qy 504 YKEHLGASIEAOLLKYKKEKDBNALSIIIONSADLADIIGSDIIKDYGKMBEEN 563
Db 512 IKEHLAIAIYESRLKKEKYNKDDKEVCKIINTFADIRDIIGTDYNDLSNRKLVGK 571
Qy 564 LNKVNKKRNEESLKIIFREKWDENKENVKWSAVLNKNETCKDYQPKQIPQFLRWF 623
Db 572 INTNSYVHRNKQNDKLFDEWVKIKDWNVSWFVKDVTCKE--DDIENIIPQPRWF 630
Qy 624 KEMGDDPCKEKEKTIYSFESKVECKKODCBNTCKNCKSYKKNWIDLKSEYKQV--- 680
Db 631 SEWGGDYQDQTKMI---ETLKVCEKXPCEDBNCKKCKNSYKEMI SKKCEYNAKQY 687
Qy 681 DKYTKDKKKWYDNIDEVKNEANVYLKESKECKDVNFDKIFNESNEYEDMCKKDE 740
Db 688 QBYQKGNMYK---SEFKSIKPEVYLKYSKCSNLNFEDEFKEELHSDYANKCTWCP 744
Qy 741 IKYL-----SEFKSIKPEVYLKYSKCSNLNFEDEFKEELHSDYANKCTWCP 750
Db 745 VKDVPISIRNEQTSQEAPEBSTEIAHRTETRTDERKNOEPANKOLKNPQQSVGENGT 804
Qy 751 -----KYKHD 755
Db 805 KDLAQEDLGGSRSEDEVTFQEGVANGHPKGBDOTLQKSDAIPNIGEPETGISTTESRHE 864
Qy 756 -----IYDIDTSDTF-----GGTPTISINANIEQQSG-KDTSNT----- 790
Db 865 EGRNKQALSTSVDPPELSDTLQLHEDTKENDKLFLESSTITPTESGSDTEETPSISEG 924
Qy 791 --GN-----SETSDSPVSHESPSDA--AINVEKLSGD----- 818
Db 925 PKGNEQKRDSDLSKISVSPENSREPETAQDTSNLLKLGKGDVDSMPKAVIGSSPNDNI 984
Qy 819 -----ESSSETGILDIN----- 831
Db 985 NVTEQGDNI SGVNSKPLSDVDRPKNBEVKEHTSNSDNVQSGGIVNMVYKELKDTLE 1044
Qy 832 DPSVTNNVNEH-----DASNTQGSVNTSDITN---GHSESLNR----- 869
Db 1045 NPSSSLDEGKAHELSEPNLSSDQMSNTPGPLDNTSEETTERISNNKYKNEREGERTL 1104
Qy 870 TTNAQDIKIGRSGNEQSDN---QENSSHSDNSGSLTI-QQVPSDNTQNTY----- 917
Db 1105 TKEYEDIVLKSHMRRESDDGLYDENSIDLSTVNDSEDAEAKMGKNDTSEHNSHQHIE 1164
Qy 918 -DSQNHRTDPNALA-----SLP-----SDDKINE-----LEGFDSDRDS 951
Db 1165 SDQCKNDMKTGDLGTHTVQNEISVPVTEIDEKLRESKESIKHABEERLSHTDIIKIN 1224
Qy 952 ENGRGDTTSNTHYR-----RTNIYSERVNSHDPIRGMANNNAHQYITQ-- 998
Db 1225 PEDRNSNTLHLKORNEENRHLTNQNTINISQERDLQKHG-----HTWNNLHGQVGSERS 1280
Qy 999 ---IENNGIIRGOESAGNSVNYKDN-----PKRSNPFSSENDHKQNIQYNSRDTKRV 1048
Db 1281 QINSHSHGNRODRGNSGVNLNRSNNNNFNINPSRYNL---YDKGLDLDIYENRDSST 1337
Qy 1049 REBIUKSKONKCNNEYSMEYCTYSDESNSSPGPCSRERKCLKCQISDYCLKYPNFYSI 1108
Db 1338 KELIKLABINKCENESVYKCDHMIHEEIPLKTCTKEKTRNLCAVSDYCMSTFTYDSE 1397
Qy 1109 EYTNCKISEIKSPYKCFKSGQS 1132
Db 1398 EYNNCTKREFDPSYCFRKEAFS 1421
```

RESULT 27

```
Q95VT1_PLAFA
ID Q95VT1_PLAFA PRELIMINARY; PRT; 1402 AA.
AC Q95VT1,
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte-binding antigen-175 (Fragment).
GN Name=BBA-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Li X.R., Chishti A.H., Oh S.S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF406762; AAK96216.1; -, mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy_binding; 2.
FT NON_TER 1
FT NON_TER 1402
SQ SEQUENCE 1402 AA; 163080 MW; A9E21BD8774E429A CRC64;

Query Match 23.5%; Score 1524; DB 2; Length 1402;
Best Local Similarity 28.2%; Pred. No. 7.6e-56;
Matches 402; Conservative 207; Mismatches 446; Indels 370; Gaps 43;

Qy 40 DISVNSPELNHNKNTIYDSYEDVNNKLINSFVENSKVKGRSLSPINN-----KTKS 94
Db 16 DVYKEKFNELDKKYGNVQXTD-----KKIFTIENK-----LDILNNSKFKRWKS 62
Qy 95 Y---DIIPPSVYNDKFNLSLSENEGNSGNTNSNPNANTSEISIGKONKQ----- 141
Db 63 YGTPDNIDKMSLIN-----KHNEEMFNNNYOSFLSTS--SLIKONKYYPINAVRVS 114
Qy 142 -YTFIQKRTH-----LPACGIRKRSIKWICHENSEKIT--VCVPRDKIOLCVANF 188
Db 115 ILSFLDRIRNGRNTSSNNEVLSNCRCKRGMKDKKNDKNDKSNVYCIPIPRRIQICVNL 174
Qy 189 LNSLETWKEFKEIFLISVNTAEKLYNKNEGDPSIFCNELRNSFSDFRSFGDDMDF 248
Db 175 SLIKTYTKETKOHFIEASKESQLLLKNDKNKYNSKFCNDLKNSFLDYGHLANGNDMF 234
Qy 249 GGNTDRVKGIVNTKPSDYKKNVEKLNKIKGEWEKNKANLNMHMIINHKNISKECAI 308
Db 235 GGYSTKAENKIQEVFKGAHGEISEHKIKNFRKEWNEFBREKLWEAMLSEHKNNIN-NCKN 293
Qy 309 IPAREPQINLWIKEMNENFLWEKRLFLNLIKDKCVENKYEACFGGCLPCSSYTSFPMK 368
Db 294 IPOBELQITQWIKWEHGEFLERDRNRSLPKSKNNNTLYEACEKECIDPCWKYRDMTIR 353
Qy 369 SKTQMEVLTMLYKKKNSGVDKNNFLNLFK-KNNKND--LDDFFKN-EKEYDDLCDCRYT 424
Db 354 SKFEWHTLSKEYETOK--VPKENAENYLLIKSENKNDKAVSLNLCDAEYSKYCDCKHT 411
Qy 425 ATTIKSPFANGPAKNDVDIASQINVDLRGFGCNYS--NNEKSWNCTGTFTNKPFGTCBP 482
Db 412 TTLVKSVLNGNDNTIKEREHEIDLDSPSKGDCDKNSVDNTNTKWECKPKYKLSKTDVCP 471
Qy 483 PRQOTLCIGRTYLLHRGHEEDYKEHLGASIEAOLLKYKKEKDBNALSIIIONSADL 542
Db 472 PRQELCLGNIDRIYDKNLLMIKEHLAIAIYESRLKKEKYNKDDKEVCKIINTKTADI 531
Qy 543 ADIIKGSDIIKDYGKMBEENLANKVKDKKRNESLSKIFREKWDENKENVKWSAVLK 602
Db 532 RDIIGTDYNDLSNRKLVGKINTNSYVHRNKQNDKLFDEWVKIKDWNVSWFVK 591
Qy 603 NKETCKDYDQKIPQFLRWFKEWGGDDFCERKEKIKIYSPFSFKVECKKQCDENTCKNK 662
Db 592 DKTCKE-DDIENIPQPRFSEWGGDYCQDKTKMI---ETLKVCEKXPCEDDNCKRKC 647
```

```
QY 663 SEYKKNIDLKSEYEKQV---DKYTKDNKNKMYNDIVKKNKEANVYLKSKCKEKKDNVF 719
D 648 NSYKEMISKKBEYNKQAKQYQYQKGNVY---SEFKSIRKPEVYLKYSKCSNLNF 704
QY 720 DDKIFNESPNEYDMCKCKDBIKYL----- 744
D 705 EDEPKELHSDYKKNKCTMCPKVDKVPISIRNNEQTSQEAVPBESIEAHRTTRTDERK 764
QY 745 -----NEIKYP----- 750
D 765 NOEPANKDLNPKQOSVGNGTKLLQBLDGLGRSEDEVTEQFGVNHGIPKGEDQTLGKSD 824
QY 751 -----KTKHD-----IYDIDTSDTF-----GDGTPISINA 776
D 825 AIPNIGEPFGISTEESRHEEGHKNQALSTVDEPELSDTQLQHDHETKNDKLPLESST 884
QY 777 NINEQSGG-KOTSN-----GN-----SETSDSPVSHPESDA--AINVEKL 815
D 885 ITSPTESSGSDTEETPSISEAPKNEQKRDSDSLKISVPENSRPETAKTNSLLKL 944
QY 816 SGD----- 820
D 945 KGDVDISMPKAVIGSSPNDNINVTQGDNISGVNSKPLSDDDVRDPKXNHEEVEKHTSNSDN 1004
QY 821 SSETRGILDIN-----DPSVTNNVEH-----DASNTQGSVNSTSDI 858
D 1005 VQSGGGVNNMVEKELKDTLENPSSSLDBGKAHEVLSEPNLSSDQDMSNTPGPLDNTSEE 1064
QY 859 TN---GHSEGSNLR-----TTNAQDIKIGRSNEOSDN---QENSSHSSDNGSLTI 904
D 1065 TTERISNNEKYNREBERLTUKYEDIVLKSHNRESDDGELYDENSLSSTVNESEDA 1124
QY 905 -GQVPSDNTQNTY-----DSQNPTRDTPNALA-----SLP-----SDDKINE-- 941
D 1125 EAKMKGNDTSMNSHQHIESDQKNDMTKVGDLGTHVQNEISVPVTGEIDELRESK 1184
QY 942 -----IEGFSRSRSENGRGDTTSTNTHVR-----RTNIVSERRVNSH 979
D 1185 ESKIHAEERLSHTDIHKINPEDRNSNTLHLKDIRNEENRHLTNQINISQERDLQKH 1244
QY 980 DFIRNGMANNNAHQYITO-----IENNGIIRGOEBSAGNSVNYKDN-----PKESNF 1027
D 1245 GF-----HTMANLHGDSVRSQINHSHGHRQDRGGNSGNVLNRSNNNNFNIPSYNL 1300
QY 1028 SSENDHKNTIQEYNSRDKRVREIILKSKONKCNNEYSMEYCTYSDERNSSFGPCSRREE 1087
D 1301 ---YDKKLDLDLYENRNDSTTKELIKKLAENKCNENISVKYCDHMTHEBIPLKCTCKE 1357
QY 1088 RKKLCCOISDYCLKYNFYSEIYVNCIKSELKSPYKCFKSEGO 1132
D 1358 TRNLCCAVSDYCMSYFTYDSEYNYCTKREPDPSYTCFRKEAPS 1402
```

## RESULT 28

```
Q9N9G9_PLARE
ID Q9N9G9_PLARE PRELIMINARY; PRT; 1433 AA.
AC Q9N9G9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Erythrocyte binding protein (Fragment).
GN Name=EBP;
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CongoLese;
RX MEDLINE=21356926; PubMed=11463469; DOI=10.1016/S0166-6851(01)00298-5;
RA Ozawa H., Kocken C.H.M., Thomas A.W.;
RT "Molecular characterization of erythrocyte binding protein of
RT Chimpanzee malaria parasite Plasmodium reichenowi.";
RL Mol. Biochem. Parasitol. 116:81-84(2001).
```

```
DR EMBL; AJ251848; CAB96159.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR InterPro; IPR008182; Pyrophosphatase.
DR Pfam; PF05424; Duffy binding; 2.
DR PROSITE; PS00387; PPAASE; UNKNOWN 1.
DR CHAIN 9 >1433 erythrocyte binding protein.
FT NON_TER 1 1
FT NON_TER 1433 1433
SQ SEQUENCE 1433 AA; 166648 MW; 733D0CDB6EDC27AE CRC64;
```

Query Match 23.4%; Score 1514.5; DB 2; Length 1433;  
Best Local Similarity 28.6%; Pred. No. 1.9e-55;  
Matches 406; Conservative 206; Mismatches 451; Indels 357; Gaps 41;

```
QY 45 NSPELNHNK-TNIYSDYEDVN-----NKLINSFVENKSVKKKSLSPINN--- 90
D 12 NEYDIKENEFLEVEYKEKFNELDKKYGNNVQTKKIFSFENK-----LDILNNSKS 64
QY 91 --KTKSY--DIIPPSYSVRNDKFNLSLENEDNSGNTSNFNANTSEISIGKDNQ--- 141
D 65 NKEWKSYGCTPDNIDKMSLIN-KYNN-----KETFNNYQSFLSTS--SLIKNRYVPIN 116
QY 142 -----YTFIQKRTH-----LFCAGIKRKSIIKWIC-RENSEKITVCVPRDKIQ 182
D 117 AVRVSRIILSFDSKINNNGRNTSSNDGLRNCREKRGNEWDCKKNGTSNVVCIPDRIQ 176
QY 183 LCVANFLNRLTMEKFEIFLISVTEAKLLNKNEGKDPSPICNELRNSPDSFSSFI 242
D 177 LCIVNLSIIKTYTKETMKDHFIEASRRSQLLKQNDNNYNSKPCNDLKNFLDYGHLM 236
QY 243 GDDMDFGGNTDRVKGVIINTKFSDYKKNVEKLNKIKKEWKEKANLWNNHMTVNHKNI 302
D 237 GNDMDFGGSTKAEINKIQEVFGVGHGKISEHEIKNFKKWNNEFREKLWEAMISEHKNL 296
QY 303 SKECALIAPAEPOINLWIKENENFLMEKKRFLNFKDCVENKRYEACFGGRLPCSSY 362
D 297 S-DCKNIPPEELQINQIKWEHGEFLSERNNRLKPKSKCNNTLYEACAKECIDPCMKY 355
QY 363 TSPMKSKQMEVLTNLYKKKN-SGVDKNPNLNDLFKNKNKNDLDDFF-KNEKEYDDLCD 420
D 356 RDWIIKSKFEWHTLSKEYERKNVSNKDAEYKLIK-FSKNDAKVSLLLDCKDAEYSKYCD 414
QY 421 CRYTATIIKSFNLGPAKNDVDIASQINVNDLGRFGCNYKS--NNEKSNWCTGTFTNKPFG 478
D 415 CKHTTLVKSVLNGKONTKEERETIDDDFSKFGCDKNSVDNTRKEWCKKPYILSTKD 474
QY 479 TCPEPPRRQTLCLGRTYLLHRGHEEDYKEHLLGASIIYEAQLLKYYKEKDNALCSIIQNS 538
D 475 VCVPPRRQELCLGNIDRIYDKNLLMIKEHILATAIYESRILKRYKKNKDDNEVCNVINKS 534
QY 539 YADLADIIKGSDDIKDYGGKMEENLNKVKDKKRNEESLKI PREKWDNDENKENVKMS 598
D 535 FADIRDIIKGSDYWNDLNRLKLVGKINTNSKYVHKKNENDKLPDEWNNVYKDDWNNVMS 594
QY 599 AVLKNKETCKDYDFOKIPQFLRWFKWEGDGFCEKKEKIYSPESPKVECKKDCDENTC 658
D 595 WVPKDKTVCKE-DDIENIQPFNFWSEWGGDDYQ---YKLMIDTLKACBEKGYDITC 650
QY 659 KKKSEYKQWIDLKSEYEKQVDKYT--KDKNKKMYDNDIVKKNKEANVYLKSKCKECKD 716
D 651 KKKCSYKQWISKQKYLKQVITYIKEYQRNNRIY---PIVKTMEPKVYLKEYSKKCSN 707
QY 717 VNFDDKIFNESPNEYEDMC-----KKCDRIKYLNEIKYP 750
D 708 INFEDFNEBVEHSDYKKNKCMISTKVLDPVIPABKNVTEALETVVPKENTIERKNESDTE 767
QY 751 KTKH-----DIYDIDTSDTFDGDGTPIS-----INANI---NEQSQS 784
D 768 EQKNQEPNNYLKDSQQRVGENGCTKYSSQEDLGRSNEEGVTQDFGLNDIIPKGEQISG 827
```

QY 785 KD-----TSNTGSETSDSPVSHE-----PES 806  
 Db 828 KSGVIONRSEMGISTSESRHEGHNEQTLSVDQSELSDTLQSHEDTKENDELPLA 887  
 QY 807 DAAINVEKLGDESSETRGILDNDPSVTNNVEHDASNTQGSVNTSDITNGHSESS 866  
 Db 888 SSTTSPMESGSSDTEETQ---SISEGPKGNEQKERRDDSSSKISVSPEISRTETHDKOT 944  
 QY 867 LNRITTAQDIK-----IGRSGNEQSDNQENSSH-SSDNGSGSLTIGQVPS- 910  
 Db 945 SNLLRLKEDVDISMPKAVIGKDPNDHINNTAEGDHSIGVNSPLSDGVRPKNHEEVKEH 1004  
 QY 911 -----DNTQNTVDS-----QNPRTD-PNALASL-----PS 935  
 Db 1005 TSNPDNVQSGENVMNVENQDKTLTENTSSSLGEGKAHELSBENLSSGQDMSTNAPL 1064  
 QY 936 DD-----KINE-----LEGFDSRDSENG----- 954  
 Db 1065 DNTSEETTERISNNEYKVNREDERTLKEYEDIVLKSHMNRSDDBELDGKNSDVPVN 1124  
 QY 955 -----RGDTTSTHVD----- 965  
 Db 1125 DESEDEVEKMRNDTSELSESSQHSIEDQENDMKTVDGLTIHVQNEINDSVTGENKE 1184  
 QY 966 -----RRTNIVSRRVNSHDFIRNGMANNAH-----H 993  
 Db 1185 QTVHDEGVRLSHDKKHINLEDKNSNLIHLKMRNEERKQLANQINISQQRDLQEH 1244  
 QY 994 QYIT--QIENNGIITRGQESAGNSVNYKONPKRS-----NFSSEN-----DHK 1034  
 Db 1245 GFHTWNNLHGVSSEGSGQINHGPNVPRQDRGKPSGNVLNMGNNNNFNIFSRNLYDKK 1304  
 QY 1035 KNIQEYNSRDKRVREIEIILSKQKNNEYSMEYCTYSDRNSSPGCSREERKLCQ 1094  
 Db 1305 LNLDLVNRNDSTYKELIKLASINKCEBISAKYCSMDIKDEIPLKCTCTKEKINLCCA 1364  
 QY 1095 ISDYCLKYFNYSIYYNCIKSEIKSPYKCFKSEGOSSI 1134  
 Db 1365 VADYCLSYFTVDSKEYSNCTKREFEDPSYTCFRKKGFSSM 1404

## RESULT 29

QY Q81796 PLAFPA PRELIMINARY; PRT; 614 AA.  
 Db Q81796;  
 AC 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Erythrocyte binding antigen region II (Fragment).  
 GN Name=eba-175;  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22588500; PubMed=12702678;  
 RA Baum J., Thomas A.W., Conway D.J.;  
 RT "Evidence for diversifying selection on erythrocyte-binding antigens  
 of Plasmodium falciparum and P. vivax."  
 RL Genetics 163:1327-1336(2003).  
 DR EMBL; AJ438810; CAD27535.1; -; Genomic DNA.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0009405; P:pathogenesis; IEA.  
 DR InterPro; IPR008602; Duffy\_binding.  
 DR Pfam; PF05424; Duffy\_binding; 1.  
 FT NON\_TER 614 614  
 SQ SEQUENCE 614 AA; 73292 MW; A168F604AB322F7F CRC64;

Query Match 18.7%; Score 1210.5; DB 2; Length 614;  
 Best Local Similarity 38.2%; Pred. NO. 4.3e-43;  
 Matches 245; Conservative 112; Mismatches 239; Indels 45; Gaps 11;

QY 151 LPACGIKRSIKWICRENSEKIT--VCVPDRKIQLCVANFLNSRLTETMEKPEIFLISVNT 209  
 Db 11 LSSCREKRGKMKDCKKNDNSYVCIPDRIQICIVNLSIIKTYTETKMDHFEASKK 70  
 QY 210 EAKLLYNNEGKQPSIFCNELNRSFSDPRSSFIGDMDFGNTDRVKGYINTKSDSYKE 269  
 Db 71 ESQALLKNDNKNYSKFCNDLKNFLDYGHLAMGNDMDFGYSYTKAENKIQEVFKGAHGK 130  
 QY 270 KNVSEKLANNIKKEWEKKNANLWNIHVKNIKECAIIPAEPOINLWIKENNENFLM 329  
 Db 131 ISEHEIKNFKREWNEFREKLWEALWSEHKNNIN--NCKNIPQELQITQWIKWHGEFLL 189  
 QY 330 EKKEFLNLNIDKQVENKYEACFCGCRLLPCSSYTSFMKSKSTQBEVLNLYKKNGSGYDK 389  
 Db 190 ERDNRSLPKSKCKNNNTLYEACEKIDPCMKYRDWIIRSKFEWHTLSKEYETQN--VSK 247  
 QY 390 NNFNDLFPKNNKNDLDDFFKN--EKEYDDLDCCRYTATIIKSLFNGPAKNDVDIASQIN 447  
 Db 248 ENAENYLIKKMMDAKVSLLLNNDCAEYSKYCDCKHTTTLVKSLNGNDNTIKEKREHID 307  
 QY 448 VNDLRGFGCNYKS--NNEKSNWCTGTFTNKPFGTCEPPROTLCIGRTYLLHRGHEEDYK 505  
 Db 308 LDDFSKFGCDKNSVDTTNKWECKPKYKLTQKVCVPPRROELCLGNIDRIYDKNLLMIK 367  
 QY 506 EHLGASIIYEAQLKYKKEKENALCSIIQNSYADLADIIGSDIIKDYVYKKEENLN 565  
 Db 368 EHILAIAIYESRLIKRYKNDKEVCKIINKTPADIRDIIGGTDYNDLNRKLVGKIN 427  
 QY 566 KVNKDKRNEESLKIPEKWDENKENVWMSAVLNKETCKDYDFKQKIPQFLRWPK 625  
 Db 428 TNSNYVRHNQNDKLPDAMWVKIKDQVNVISWVFKDKTVCKE--DDIENIPQFPRWFE 486  
 QY 626 WGDPPCEKREKIKYSPFSFVCKKCDKDCENTCKNCKSEYKKWIDLKSEYKQV---DK 682  
 Db 487 WGDYCYQDKTKMI---ETLKVCKECPCEDNCKSKNSYKEMISKKKEETNKAQKQOE 543  
 QY 683 YTKDNKKMVDNIDEVKNKEANVYLKESKCKDQVNFDDKIFNESPNEBYEDMKCKCDRIK 742  
 Db 544 YQKGNVYKMY---SEPKSIKPEVYLKYSKKCSNLNFEDEPKELHSDYKKNKCTCPVK 600  
 QY 743 YLNEIKYPTKHDYIDITFSDTFGDTGTPISINANINEQOS 783  
 Db 601 -----DVPISIRN-NEQTS 614

## RESULT 30

QY Q81791 PLAFPA PRELIMINARY; PRT; 616 AA.  
 Db Q81791;  
 AC 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Erythrocyte binding antigen region II (Fragment).  
 GN Name=eba-175;  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22588500; PubMed=12702678;  
 RA Baum J., Thomas A.W., Conway D.J.;  
 RT "Evidence for diversifying selection on erythrocyte-binding antigens  
 of Plasmodium falciparum and P. vivax."  
 RL Genetics 163:1327-1336(2003).  
 DR EMBL; AJ438822; CAD27547.1; -; Genomic DNA.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0009405; P:pathogenesis; IEA.  
 DR InterPro; IPR008602; Duffy\_binding.  
 DR Pfam; PF05424; Duffy\_binding; 1.  
 FT NON\_TER 616 616  
 SQ SEQUENCE 616 AA; 73292 MW; A168F604AB322F7F CRC64;

```
SQ SEQUENCE 616 AA; 73559 MW; 4B3BE72152B7D1F0 CRC64;

Query Match 18.7%; Score 1209.5; DB 2; Length 616;
Best Local Similarity 37.6%; Pred. No. 4.7e-43;
Matches 253; Conservative 115; Mismatches 236; Indels 69; Gaps 15;

QY 121 NTNSNNPANTSEISIGKDNKQYTFIQKRTHLFACGIGKRSIKWICRENSKIT-VCPVDR 179
Db 3 NTSSNN-----EV-----LSSCREKRGKMKWDCKKNDNSNYVCIPDR 40

QY 180 KIQLCVANFLNSRLTMEKFEIPLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFS 239
Db 41 RIQLCIVNLSIIKTYTKETMKDHFIEASKESQLLLKQNDNYSKFCNDLKNISFLDYGH 100

QY 240 SFTGDDMDFGNTDRVKGINTKPSDYKKEKNVEKLNKIKKEWKEKNKAMLNHMIYNHK 299
Db 101 LAMGNDMDFGYSTKAENKIQEVFGAHSIEHKIKNFRKEWNEPREKLWEAMLSHK 160

QY 300 GNISKECALIPABEPOINLWIKENENFLMEKRLFLNIDKQCNVKKYBACFGCRLPC 359
Db 161 NNIN-NCKNIPQBELQITQWIKEMHGFLLERNRSKLPKSKCKNNTLYACEKECIDPC 219

QY 360 SSVTSFMKSKTOMEVLTNLYKKNSGVQKNNFLNDFK-KNNKND--LDDFFKN-EKEY 415
Db 220 MKYRDMIIIRSKFEWHLSKEYETQK--VPKENAENYLIKISEKNQDAKVSLLLNNDCAEY 277

QY 416 DDLCDCRYTATIIKSFNGFLNGPAKNDVDIASQINVDNLRGFCNYSK--NNEKSNCTGTF 473
Db 278 SKYCDCKHHTTLVKSVLNGNDNTIKEREHIDLDDFSFGCDKNSVDNTNTKWECKNPIY 337

QY 474 NKPPGTCPPRRQTLCLGRITLHRRGHEEDYKEHLGASIEAQLKLYKKEKDNALCS 533
Db 338 LSTKDVCPVPRROELCGNIDRIYDKNLLMIKEHILAIAIESRLKRYKNKDDKEVCK 457

QY 594 WKMSAVLNKNETCKDYDFQKIPQFLRWPKWGGDDFCERKEKIEYSFESFKVECKKDC 653
Db 458 WNVISWFKDKTVCKE-DDIENIPQFRWPFSEWGGDYCDQTKMI---ETLKVECKECPK 513

QY 654 DENTCKNKCSEYKWKIDLLKKSEYEQV---DKYTKDNKKMYDNIDEVQKNEANVYLKEK 710
Db 514 EDDNCKSKNSYKEWISKKKEEYNKQAKQYQYQKGNKYKMY---SEFKSIKPEVYLKY 570

QY 711 SKECKDVNFDDKIFNESPNEDYDMCKCKDEIKYLNEIKYPTKHDIDYDIDTFSDFGDT 770
Db 571 SEKCSNLNFEDEFKEELHSDYKNKCTWCPEVK-----DV 604

QY 771 PISINANINEQQS 783
Db 605 PISIIRN-NEQTS 616

RESULT 31
Q81797_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q81797;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 23, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22589500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
```

```
RL Genetics 163:1327-1336(2003).
DR EMBL; AJ43808; CAD27533.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duff binding.
DR Pfam; PF05424; Duff binding; 1.
FT NON_TER 1
FT NON_TER 616
SQ SEQUENCE 616 AA; 73589 MW; 312B7CF98D9796BE CRC64;

Query Match 18.7%; Score 1209.5; DB 2; Length 616;
Best Local Similarity 37.7%; Pred. No. 4.7e-43;
Matches 254; Conservative 112; Mismatches 238; Indels 69; Gaps 15;

QY 121 NTNSNNPANTSEISIGKDNKQYTFIQKRTHLFACGIGKRSIKWICRENSKIT-VCPVDR 179
Db 3 NTSSNN-----EV-----LSSCREKRGKMKWDCKKNDNSNYVCIPDR 40

QY 180 KIQLCVANFLNSRLTMEKFEIPLISVNTAEKLLYNKNEGKPSIFCNELRNSFSDFS 239
Db 41 RIQLCIVNLSIIKTYTKETMKDHFIEASKESQLLLKQNDNYSKFCNDLKNISFLDYGH 100

QY 240 SFTGDDMDFGNTDRVKGINTKPSDYKKEKNVEKLNKIKKEWKEKNKAMLNHMIYNHK 299
Db 101 LAMGNDMDFGYSTKAENKIQEVFGAHSIEHKIKNFRKEWNEPREKLWEAMLSHK 160

QY 300 GNISKECALIPABEPOINLWIKENENFLMEKRLFLNIDKQCNVKKYBACFGCRLPC 359
Db 161 NNIN-NCKNIPQBELQITQWIKEMHGFLLERNRSKLPKSKCKNNTLYACEKECIDPC 219

QY 360 SSVTSFMKSKTOMEVLTNLYKKNSGVQKNNFLNDFK-KNNKND--LDDFFKN-EKEY 415
Db 220 MKYRDMIIIRSKFEWHLSKEYETQK--VSKENAENYLIKISEKNDAKVSLLLNNDCAEY 277

QY 416 DDLCDCRYTATIIKSFNGFLNGPAKNDVDIASQINVDNLRGFCNYSK--NNEKSNCTGTF 473
Db 278 SKYCDCKHHTTLVKSVLNGNDNTIKEREHIDLDDFSFGCDKNSVDNTNTKWECKPYK 337

QY 474 NKPPGTCPPRRQTLCLGRITLHRRGHEEDYKEHLGASIEAQLKLYKKEKDNALCS 533
Db 338 LSTKDVCPVPRROELCGNIDRIYDKNLLMIKEHILAIAIESRLKRYKNKDDKEVCK 397

QY 534 IIONSVADLADIIGSDIIDKYGKQMEENLNKKNKKNESLKFREKKNWDEKENV 593
Db 398 IINKTFADIRDIIGGTDYNDLSNRKLGVKINTNSYVHRNKQNDKLFREDWVKVIKDV 457

QY 594 WKMSAVLNKNETCKDYDFQKIPQFLRWPKWGGDDFCERKEKIEYSFESFKVECKKDC 653
Db 458 WNVISWFKDKTVCKE-DDIENIPQFRWPFSEWGGDYCDQTKMI---ETLKVECKECPK 513

QY 654 DENTCKNKCSEYKWKIDLLKKSEYEQV---DKYTKDNKKMYDNIDEVQKNEANVYLKEK 710
Db 514 EDDNCKSKNSYKEWISKKKEEYNKQAKQYQYQKGNKYKMY---SEFKSIKPEVYLKY 570

QY 711 SKECKDVNFDDKIFNESPNEDYDMCKCKDEIKYLNEIKYPTKHDIDYDIDTFSDFGDT 770
Db 571 SKRCSNLNFEDEFKEELHSDYKNKCTWCPEVK-----DV 604

QY 771 PISINANINEQQS 783
Db 605 PISIIRN-NEQTS 616

RESULT 32
Q817A0_PLAFA PRELIMINARY; PRT; 614 AA.
AC Q817A0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-175;
```



```

Db 514 EDDNCKSKNSYKEMWISKKEEYNKQAKQYQYQKGNYYKMY---SEFKSIKPEVYLKXY 570
Qy 711 SKECKDVNFDKIPNESPNYEYDMCKDEIKYLNELKYPKTKHDIYDIDTFSDTFDGGT 770
Db 571 SEKSNLNFDEFKELHSDYKQKNTCTMCPVK-----DV 604
Qy 771 PISINANINEQOS 783
Db 605 PISIIRN-NEQTS 616

RESULT 34
Q25739_PLAFA
ID Q25739_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q25739;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=EBA-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=VI-9;
RX MEDLINE=97237565; PubMed=9084043; DOI=10.1016/S0166-6851(96)02791-0;
RA Liang H., Sim B.K.;
RT "Conservation of structure and function of the erythrocyte-binding
RL Mol. Biochem. Parasitol. 84:241-245(1997).
DR EMBL; U27396; AAB51620.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy_binding.
DR Pfam; PF05424; Duffy_binding; 1.
FT NON_TER 1 616
FT NON_TER 616
SQ SEQUENCE 616 AA; 73517 MW; 2C1B4990CF732693 CRC64;

Query Match 18.6%; Score 1206.5; DB 2; Length 616;
Best Local Similarity 37.7%; Pred. No. 6.3e-43;
Matches 255; Conservative 115; Mismatches 231; Indels 75; Gaps 17;

Qy 121 NTSNNFANTSBEISIGKDNKQYTFIQKRTLHACGIRKSIKMTCRENSEKIT--VCVPDR 179
Db 3 NTSNN-----EV-----LSNCRKRGKMGKWDCKKKDNRSNYVCIPDR 40

Qy 180 KIQLCVANFLNSRLTWEKPEIFLISVNTVEAKLLYNKNGKDPISIFCNELRNSFSDPRS 239
Db 41 RIQLCVNLNSIIKTYTETKWDHFIASKESQLLLKKNDKNKNSKFCNDLKNFLDYGH 100

Qy 240 SFIGDDDFGNTDRVGYINTKFSYVYKGNVEKLNIIKKEWKKKKNLWNHMIYNHK 299
Db 101 LAMGNDMDFGYSYKAENKQEVFGAGHELSEHEIKNFRKKNWNEPREKLWEAMUSEHK 160

Qy 300 GNISKECAIIPAEPPQINLMIKEWNEFLMEKRLFLNIDKCVENKYEACFGGCRLLPC 359
Db 161 NNIN-NCKNIPQEELOITQMIKEWHGFEFLERDNRSLPKSKCKNNTLYEACEKCIDPC 219

Qy 360 SSYTSFMKSKTQMEVLTNLYKKNSGVDRKNFLNDLPK-KNNQND--LDDFFKN-EKEY 415
Db 220 MKYRDMTIRSKPEWHLLTSKEYETQN--VSKENAENYLIKSKKNKNDKAVSLLLNCDAY 277

Qy 416 DDLCDCRYATIIKSLFNGPAKNDVDIASQ---INVNDLRGFGCNYKS--NNEKSNWCTG 470
Db 278 SKYCDCKHTTTLVKSVLNG---NNTYKEEREHLDLDDFSKFGCKNSVDNTNWKVECKN 334

Qy 471 TFTNKFPGCTPPRRQTLCTGRYTLHLRGHBEDYKEHLLGASIEYAQLLYKYKEKEDENA 530
Db 335 PYLSTKDVCPVPRRQELCLGNIDRIYDKNLLMIKEHILALAIYESILRKGYKNKDDKE 394

```

```

Qy 531 LCSIIQNSVADLADIIKSGDIIKDYGYKKQVEENLNKVNKKKNEESLKIIFREKWDENK 590
Db 395 VCKIINKTADIRDIIGGTYWDLNSRKLGVKINTNSNVYHRNKKNDKLFREDWKKVIK 454
Qy 591 ENWKVMSAVLNKTKCDYDFQKIPQFLRWPFKWGGDDFCCEKREKIKYSFSEFKVECKK 650
Db 455 KDVWNVISVFKDKTVCKE-DDIENIPQFRWFSEWGGDYCY---DKTKTIELTKVECKE 510
Qy 651 KDCDENTCKNKSYYKKWIDLKKSEYKQV---DKYTKDKNKKQYDNIDEVKNEANVYL 707
Db 511 KPCEDDNCKSKNSYKEMI SKKEEYNKQAKQYQYQKGNYYKMY---SEFKSIKPEVYL 567
Qy 708 KEKSKCKOVNPDKIFNESPNEYEDMCKCKDEIKYLNELKYPKTKHDIYDIDTFSTFG 767
Db 568 KKYSEKSNLNFDEFKELHSDYKQKNTCTMCPVK----- 602
Qy 768 DGTPIISINANINEQOS 783
Db 603 -DVPISIIRN-NEQTS 616

RESULT 35
Q25736_PLAFA
ID Q25736_PLAFA PRELIMINARY; PRT; 614 AA.
AC Q25736;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=EBA-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PNGL;
RX MEDLINE=97237565; PubMed=9084043; DOI=10.1016/S0166-6851(96)02791-0;
RA Liang H., Sim B.K.;
RT "Conservation of structure and function of the erythrocyte-binding
RL Mol. Biochem. Parasitol. 84:241-245(1997).
DR EMBL; U27393; AAB51617.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy_binding.
DR Pfam; PF05424; Duffy_binding; 1.
FT NON_TER 1 614
FT NON_TER 614
SQ SEQUENCE 614 AA; 73334 MW; B1994894CE1902C3 CRC64;

Query Match 18.6%; Score 1204.5; DB 2; Length 614;
Best Local Similarity 38.2%; Pred. No. 7.7e-43;
Matches 244; Conservative 110; Mismatches 239; Indels 45; Gaps 11;

Qy 154 CGIKRKSIIKWCENSEKIT--VCVPDRKIQLCVANFLNSRLTWEKPEIFLISVNTVEAK 212
Db 14 CREKRGKMGKWDCKKKDNRSNYVCIPDRRIQLCVNLNSIIKTYTETKWDHFIASKESQ 73

Qy 213 LLYNKGKDPISIFCNELRNSFSDPRSFTGGDDMGNTDRVKGYNITNKFSDYVYKKNV 272
Db 74 LLLKNDKNKNSKFCNDLKNFLDYGHLAGNDMGDFGYSYKAENKIQEVFGAGHELSE 133

Qy 273 EKLNNIKCWWEKKNLWNHMIYNHKNISKECAIIPAEPPQINLMIKEWNEFLMEKK 332
Db 134 HEIKNFRKKNWNEPREKLWEAMUSEHKNNIN-NCKNIPQEELOITQMIKEWHGFEFLERD 192

Qy 333 RLFLNIDKCVENKYEACFGGCRLLPCSSYTSFMKSKTQMEVLTNLYKKNSGVDRKNF 392
Db 193 NRSKLPKSKCKNNTLYEACEKCIDPCMKYRDWIIRSKPEWHLLTSKEYETQN--VSKENA 250

Qy 393 LNDLFKKNKNDLDDFFKN--EKEYDDLCDCRYATIIKSLFNGPAKNDVDIASQINVD 450

```

```
Db 251 ENYLKKKQNDKAVSLLLNNCDASYSKYCDCKHTTLLVKSVLNGNDNTIKREHIDLD 310
Qy 451 LRGFGCNYS--NNEKSNWCTGTNTKPPGCEPRQTLCLGRYTLHLRHEEDYKHL 508
Db 311 FSKFGCDKNSVDYTKWCKPKYKLSYKDVCPVPRQELCLGNIDRIYDKNLLMIKHI 370
Qy 509 LGASIYEAQLKYKYKEDENALCSIIIONSVADLIIGSDIIKDYGGKQKQENLNKNV 568
Db 371 LAIAIYERILKRYKKNKDKVECKIINKTPADIRDIIGGTDYNDLSNRKLVGKINTS 430
Qy 569 KDKRNEESLKI PREKWDENKENWVMSAVLNKTKETCKDYDFQKIPQLRWFPEWGD 628
Db 431 NYVRHNKNDKLFKRDAMWVKIKDQVNVISWVFKDKTVCKE--DIENIPQFRWPFSE 489
Qy 629 DFCRKRKEIYSFESFKVECKKDCDENTCNKCSYKKWIDLKKSSEYKQV---DKYTK 685
Db 490 DYCDQKTKMI---ETLKVECKEKECEDNCKSKNSYKESWISKKKEEYNKQAKQYEQK 546
Qy 686 DNKKQMDYNDIDEVKNKEANVYLKESKCEKDVNFDDKI FNSPNEYEDMCKCKDEIKYLN 745
Db 547 GNNYKMY---SEFKSIKPEVLYKYSKCSNLNFEDEBPKBELHSDYKKNKCTMCPYK 600
Qy 746 EIKYPTKTHDIYDITFSDTGDGTPISINANINEQOS 783
Db 601 -----DVPISIRN-NEQTS 614

RESULT 36
Q8I790_PLAFA PRELIMINARY; PRT; 614 AA.
AC Q8I790;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
PL Genetics 163.1327-1336(2003).
DR EMBL; AJ38824; CAD27549.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy_binding; 1.
FT NON TER 614
SQ SEQUENCE 614 AA; 73306 MW; E7E6CD3C48774839 CRC64;

Query Match 18.6%; Score 1204.5; DB 2; Length 614;
Best Local Similarity 38.1%; Pred. No. 7.7e-43;
Matches 244; Conservative 112; Mismatches 240; Indels 45; Gaps 11;

Qy 151 LPACGIRKRSIKWICRENSKIT--VCVPDRKIQLCVANFLNSRLTWEKPKFIPLISVNT 209
Db 11 LSSCREKRGKMKWCKKQNDRSNYVCIPDRIQLCIVNLSIIKTYTETMEDHIEASKK 70
Qy 210 EAKLLYNQNEKDPISIFCNELNRSFSDPRSSIFGDMDFGNTDRVKGYINTKFSDDYKE 269
Db 71 ESQILLKKNKNSKFCNDLKNFLDYGHLANGNDMDFGYSTKAENKIQEVFKGAHK 130
Qy 270 KNVEKLANIKKEWKEKQKANTWNHIVNHKNIKSEKAIIPAESEPOINLWIKENENFLM 329
Db 131 ISEHEIKNFRKWNNEFREKLWEAMLSEHKNNIN--NCKNIPOBELQITQWIKWHGEFLL 189
Qy 330 EKRLFLNIKDKVENKKYKACFGCRLPCSSYTSFMKSKSTQMEVLTNLYKKQNSGVDK 389
```

```
Db 190 ERDNRSKLPKSKQNTLYEACEKECDPCMKYRDMWIRSKFEWHITLSKEYETQN--VSK 247
Qy 390 NNFNDLFLKQNNKNDLDDFFQN--EKSYDDLDCRYTATIIKSPFNGPAKNDVDIASQIN 447
Db 248 ENAENYLIKKQNDKAVSLLLNNCDASYSKYCDCKHTTLLVKSVLNGNDNTIKREHID 307
Qy 448 VNDLRGFGCNYS--NNEKSNWCTGTNTKPPGCEPRQTLCLGRYTLHLRHEEDYK 505
Db 308 LDDFSKFGCDKNSVDYTKWCKPKYKLSYKDVCPVPRQELCLGNIDRIYDKNLLMIK 367
Qy 506 EHLGASIYEAQLKYKYKEDENALCSIIIONSVADLIIGSDIIKDYGGKQKQENLN 565
Db 368 EHILAIYERILKRYKKNKDKVECKIINKTPADIRDIIGGTDYNDLSNRKLVGKIN 427
Qy 566 KVNKDKRNEESLKI PREKWDENKENWVMSAVLNKTKETCKDYDFQKIPQLRWFPEK 625
Db 428 TNSNYVRHNKNDKLFKRDAMWVKIKDQVNVISWVFKDKTVCKE--DIENIPQFRWPFSE 486
Qy 626 WGDGFCRKRKEIYSFESFKVECKKDCDENTCNKCSYKKWIDLKKSSEYKQV---DK 682
Db 487 WGDYCDQKTKMI---ETLKVECKEKECEDNCKSKNSYKESWISKKKEEYNKQAKQYQ 543
Qy 683 YTKDNKKQMDYNDIDEVKNKEANVYLKESKCEKDVNFDDKI FNSPNEYEDMCKCKDEIK 742
Db 544 YQKGNMY---SEFKSIKPEVLYKYSKCSNLNFEDEBPKBELHSDYKKNKCTMCPYK 600
Qy 743 YLNEIKYPTKTHDIYDITFSDTGDGTPISINANINEQOS 783
Db 601 -----DVPISIRN-NEQTS 614

RESULT 37
Q8I0C8_PLAFA PRELIMINARY; PRT; 616 AA.
AC Q8I0C8;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Erythrocyte binding antigen region II (Fragment).
GN Name=eba-175;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22588500; PubMed=12702678;
RA Baum J., Thomas A.W., Conway D.J.;
RT "Evidence for diversifying selection on erythrocyte-binding antigens
of Plasmodium falciparum and P. vivax.";
PL Genetics 163.1327-1336(2003).
DR EMBL; AJ38814; CAD27539.1; -; Genomic DNA.
DR EMBL; AJ38820; CAD27545.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy_binding; 1.
FT NON TER 616
SQ SEQUENCE 616 AA; A0347CAFA5214D47 CRC64;

Query Match 18.6%; Score 1204.5; DB 2; Length 616;
Best Local Similarity 37.6%; Pred. No. 7.7e-43;
Matches 253; Conservative 113; Mismatches 238; Indels 69; Gaps 15;

Qy 121 NTNSNPNANTSEISIGKNDKQYTFIQKRTHLFACGIRKRSIKWICRENSKIT--VCVPDR 179
Db 3 NTSSNN-----EV-----LSNCREKRGKMKWCKKQNDRSNYVCIPDR 40
Qy 180 KIQLCVANFLNSRLTWEKPKFIPLISVNTAEKLLYNKNGKQDPSIFCNELNRSFSDPR 239
Db 41 RIQLCIVNLSIIKTYTETMKDHPFEASKESQLLKKNDKNKNSKFCNDLKNFLDYG 100
```





Db 338 LSTKDVCPPRRQELCLGNIDRIYDKNLLMIKEHILAIAIYESRILKRYKNKDDKEVCK 397

Qy 534 IIQNSYADLADIIGSDIIKDYYGKMEENLNKVNKDKKRNESLKIIFREKMWDEKNENY 593

Db 398 IINKTFADIRDIIGGTDYWNDSLNRKLVGKINTNSKYVHRNKKNDKLFDEWVKVIKKDV 457

Qy 594 WKUMSAVLKNKCKDYDKFKIQFLRWPKWEGDDFCEKKEKIKIYSPESFKVECKKDC 653

Db 458 WNVISWVFKDRTVCKE-DDIENIPQFRWFSEWGDDYQDKTKMI---ETLKVECKEPC 513

Qy 654 DENTCKNKCSEYKKWIDLAKSEYEKQV---DKYTKDNKKMYDNIDEVKNKEANVYLKEK 710

Db 514 EDDNCKSKCNSYKEWISKKKEEYNKQAKQOEYQKGNVYKMY---SEFKSIKPEVYLKY 570

Qy 711 SKECKDVNFDDKIFNESPNEDMCKKDEIKYLNIEIKYPTKHDIYDIDTFSTFGDGT 770

Db 571 SEKCSNLNFEDEFKEELHSDYKNKCTMCPEVK-----DV 604

Qy 771 PISINANINEQOS 783

Db 605 PISIIRN-NEOTS 616

Search completed: November 21, 2005, 20:38:44  
Job time : 274 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:26:57 ; Search time 51 Seconds  
(without alignments)  
2282.790 Million cell updates/sec

Title: US-10-677-980-2

Perfect score: 6481

Sequence: 1 MKGYFNIYFLIFLYNVI.....VQETNISDYSEYNNKXMY 1210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1557	24.0	1435	2 A37793	erythrocyte-bindin
2	805.5	12.4	1045	2 T18373	erythrocyte bindin
3	765.5	11.8	1070	2 T30848	Duffy receptor - P
4	723.5	11.2	1153	2 T28652	erythrocyte bindin
5	655	10.1	778	2 A35970	erythrocyte-bindin
6	623.5	9.5	2706	2 T28155	variant-specific s
7	600	9.3	3006	2 T28625	variant-specific s
8	555.5	8.6	2924	2 T18378	variant-specific s
9	545.5	8.4	3026	2 T28431	variant surface pr
10	498.5	7.7	3078	2 T28432	variant-specific s
11	483	7.5	2212	2 T28157	erythrocyte membra
12	456	7.0	1701	2 T09127	probable erythrocy
13	438	6.8	1711	2 C71625	variant-specific s
14	433.5	6.7	2664	2 T28636	variant-specific s
15	410.5	6.3	2485	1 H71621	serine/threonine-s
16	409.5	6.3	2647	2 T28161	variant-specific s
17	401	6.2	3724	2 T18427	hypothetical prote
18	397.5	6.1	2135	2 T14602	variant-specific s
19	395	6.1	2523	2 T18477	hypothetical prote
20	392	6.0	2197	2 B71600	variant-specific s
21	379.5	5.9	1729	2 T18396	erythrocyte membra
22	370.5	5.7	1979	2 C71622	hypothetical prote
23	370.5	5.7	2228	2 T14029	variant-specific s
24	366.5	5.7	1127	2 T28317	ORF MSV156 hypothe
25	362.5	5.6	1516	2 E71619	RAD2 endonuclease
26	359.5	5.5	2401	2 T28676	rhopty protein -
27	358	5.5	1308	2 E71622	probable membrane
28	357	5.5	3844	2 T18402	asparagine/asparta
29	356.5	5.5	3394	2 T18501	hypothetical prote

ALIGNMENTS

RESULT 1

A37793

erythrocyte-binding antigen 175 - malaria parasite (Plasmodium falciparum)

C;Species: Plasmodium falciparum

C;Date: 30-Apr-1991 #sequence\_revision 30-Apr-1991 #text\_change 09-Jul-2004

C;Accession: A37793; S11561

R;Sim, B.K.L.; Orlandi, P.A.; Haynes, J.D.; Klotz, F.W.; Carter, J.M.; Camus, D.; Zegans

J. Cell Biol. 111, 1877-1884, 1990

A;Title: Primary structure of the 175K Plasmodium falciparum erythrocyte binding antigen

A;Reference number: A37793; MUID:91035602; PMID:2229177

A;Accession: A37793

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1435 <SIM>

R;Sim, B.K.L.

Mol. Biochem. Parasitol. 41, 293-295, 1990

A;Title: Sequence conservation of a functional domain of erythrocyte binding antigen 175

A;Reference number: S11561; MUID:90377299; PMID:2204835

A;Accession: S11561

A;Molecule type: DNA

A;Residues: 995-1158 <SIW>

A;Cross-references: UNIPARC:UPI00001785PB; EMBL:X52524

Query Match 24.0%; Score 1557; DB 2; Length 1435;

Best Local Similarity 28.5%; Pred. No. 4.5e-59;

Matches 410; Conservative 214; Mismatches 467; Indels 348; Gaps 42;

Qy	1	MKGYNFYELIPLIFLYNVRINESIIIGRTLYNRQDESSDISRVNSPELNHHKNTIYDS	60
Db	1	MKNISYIFASFFVLFAKARNEYDI-----KENEFKLDVYKEKFNELDKKYGKQKT	55
Qy	61	DYEDVNNKLINSFVENKSVKKGSLSPINN-----KTKSY---DIIPPSYRNDRKFNLSL	112
Db	56	D-----KKIFTRIENK-----LDILNKKFNKRWKSYGTPDNIDKNLSLIN-----	96
Qy	113	SEHEDNGNTNSNINFANTSISIGKONKQ-----YTFIQKTH-----	150
Db	97	KHNEEFNNYQSFLSTS--SLIKQNKYPINAVRVSRILSFLDSRINNGRNTSSNEV	154
Qy	151	LFAGGIKKIKIKICRENSKIT--VCVPDRKIOLCVANFLNSRLETWEKKEIFELISVNT	209
Db	155	LSNCREKRGKMDCKKNDKRSNVCIPDRRIQICIVNLIIKIITYTKETWKDHFIEASKK	214
Qy	210	EAKLLYNKNEGKPSIFCNELRNSFSDRSFSGDDMDFGNTDRVKGYNITKFSDYKKE	269
Db	215	ESQLLKKNDKYNKSKFCNDLKNISFLDYGHLAGNDMDFGYSTKKAENKIQEVFKGAHGE	274
Qy	270	KNVEKLANIKKEWKEKKNALNNHMIIVNHKGNISKECAIIPAESFQNLWIKENWENFLM	329
Db	275	ISEHKIKNFKREWNNEFKLEWAMLSEHKNNIN--NCKNIPQEBELQITOWIKEHGEPFL	333

```
QY 330 EKKELFLNIKDKVENKYYACFGGCLPCSSYTSFMKKSKTQMEVLTNLYKKNSGVDK 389
Db 334 ERNRSKLPKSKCKNNTLYEACEKICDPCMKYRDWTIIRSKFEWHITLSKEYETQK--VPK 391
QY 390 NNFLNDLFFK--KNKNKD--LDQFFKN--EKEYDDLCDCRYTATIIFKSPFLNGPAKNDVDIASQ 445
Db 392 ENAENYLIKISEKNDKAKVSLLLNNCDABYSKYCDCKHTTLYKSVLNGNDNTIKEKREH 451
QY 446 INVNDLFGGNGYKS--NNEKSNWCTGTFNKKFGTCPPRRQTLCLGRYTLHLRGHEED 503
Db 452 IDLDDFKFGCDKNSVDNTNKWECKNPFYLLSTKQVCPPRQELCLGNIDRIYDKNLLM 511
QY 504 YKEHLGASIEAQLLYKKEKDNALCSIIQNSYADLADIIGSDIIKDYKQKQKEEN 563
Db 512 IKEHILAIATYESILKRYKNKDDKEVKCIINKTFADIRDIIGDTYWNDSLNRKLGVK 571
QY 564 LNKVNKDKKNEESLKIFREKMWDEKENYKVMASAVLNKKECTCKDYDKFKQIPQFLRWF 623
Db 572 INTNSKYVHRKNDKLFREHWKVIKQVNVVSWFKTVCKE--DDIENIPQFFRWF 630
QY 624 KEWGDDFCERKKEKIYFSGFKVECKKCDENTCKNCKSEYKWKIDLKSEYEQV--- 680
Db 631 SEWGDDYQDQTKMI---ETLKVECKEPCEDDCKCKSKNSYKESWISKKEEYNKQAKQY 687
QY 681 DKYTKDNKQKMYDNIDEVKNEANVYLKESKECKDVNFDKIFNESPNEYEDCKKQDE 740
Db 688 QEYQKGNKYMY---SEFKSIKPEVLYLKYSKCSNLNFEDEFKBLHSDYKKNCTWCPE 744
QY 741 IKVL-----NK----- 746
Db 745 VKQVPISIIIRNEQTSQEAPEENTEIAHRTETPSISEGPKGNEQKERRDDSLSKISVSP 804
QY 747 -----IKPKT----- 752
Db 805 ENSRPETDAKOTSNLLKLGVDISMPKAVIGSSPNDINNVTEQDNIIGVNSKPLSDDV 864
QY 753 ---KHDYDID-----TFSDTFDGTGPIST----- 774
Db 865 RPKCKELEDQNSDESEETVNVNHSKSPSINNDDSGSGSATVSESSSNTGLSIDDDRNG 924
QY 775 -----NANINEQSGKOTSNGTSETSDSP---VSHEPESDAAINV 812
Db 925 DTFVRTQDANTEVDIRKENADQDEKGADEERHSTSELSPEEKMLDNEGGSNLNH 984
QY 813 EKL-----SGDESSETGILDIN-----DPSTVNNVNEVH-----DAS 846
Db 985 EEVKEHTSNSDNVQSGGIVMNVYKELKDTLENPSSSLDEGRAHELSBPNLSSDQDMS 1044
QY 847 NTQGSVNTSDITN---GHSESSLNR-----TTNAQDIKIGRSGNEQSDN-----QENS 892
Db 1045 NTPGLDNTSEETTERISNNEYKVNREDEBRTLTKYEYEDIVLXSHMNRSDGDELVDENS 1104
QY 893 SHSNDNSGLTI--QGVPSDNTQNTY-----DSQNPHRDTPNALA-----SLP- 934
Db 1105 DLSIVNDESEDAEAKMGKNTSESHNSQHIESDQKQDMKVTGDLGTHVQNEISVPV 1164
QY 935 ---SDDKINE-----IEGPDSSRDSNGRGDTTSNTHDVR-----R 967
Db 1165 TGEIDEKLSKESKIHKAEBERLSHTDIHKINPEDRNSNTLHLKOIRNEENERHLTNQN 1224
QY 968 TNIVSERVNSHDFIRNGMANNAHQYITQ-----IENNGIIRGQESAGSNVYKDN- 1021
Db 1225 INISQERDLQHGPF---HTWNLLHGGVGSERSQINSHHGHRQDRGNSGNYLNRSSN 1280
QY 1022 -----PKRSNFSSENHKKNIQYNSRSDTKRYREEIIKLSKONKCNNEYSMEYCTYSDE 1075
Db 1281 NNFNNIFSRNLY---YDKLDDLDIYENRDSSTTKELLKLAELNKENEISVYCDMH 1337
QY 1076 RNSSPGPCSBEERKLCQIISDYCLKYFNFYISIEYNCIKSEIKSPYKFKSEGOSSI 1134
Db 1338 EEIPLKTCCTEKRNLNCCAVSDYCMYSFTYDSEBYNCTCKREFDDPSYTCFRKEAFSSM 1396
```

## RESULT 2

```
T18373
erythrocyte binding protein - Plasmodium cynomolgi
C;Species: Plasmodium cynomolgi
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18373
R;Okenu, D.M.; Malhotra, P.; Lalitha, P.V.; Chitnis, C.E.; Chauhan, V.S.
Mol. Biochem. Parasitol. 89, 301-306, 1997
A;Title: Cloning and sequence analysis of a gene encoding an erythrocyte binding protein
A;Reference number: Z18923; MUID:98030256; PMID:9364974
A;Accession: T18373
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1045 <OK>
A;Cross-references: UNIPROT:O15852; UNIPARC:UPI000007F49A; EMBL:Y11396; NID:e1042651; P:
C;Genetics:
A;Gene: EBP
A;Introns: 19/3; 983/3; 1008/3; 1032/3

Query Match 12.4%; Score 805.5; DB 2; Length 1045;
Best Local Similarity 26.2%; Pred. No. 3.4e-27;
Matches 269; Conservative 162; Mismatches 373; Indels 223; Gaps 40;

QY 312 EEPQINLWIKENWENFLMEKKRLFLNIKDK--CVENK--KYEACFGGCLPCSSYTSFMKK 368
Db 56 EEDNIERWLQGTNE-----RRNEENIKYKDYVTEQIKY-----AQMNEKRTSRLKK 103
QY 369 SKTQMEVLTNLYKKNKSGVDKQNNFLNDFKNNKNDLDDPFKNEKEYDDLDCRYTATII 428
Db 104 SISQINRDNVREEKQGEHKTDKTDNWERCNVLVMDYDTSNGHP-----AWTLDNV 158
QY 429 KSFL-----NGPAKNDVD---IASQ-INVNDLRGF---GCNYKSNNE-KSNM 467
Db 159 LEFVAEAGEHSLANSKSGGPNYDIDHKRTISSDVINHAFLQNVYMKKCNDRKRIRDMW 218
QY 468 CTGTFNKKFGTCPPRRQTLCLGRYTLHLRGHEEDYKEHLGASIY-----EA 516
Db 219 CP---TKK--DVCIPDRRYQLCKMEITNLVDNTNTKPHSDITPRKLYLKIYDVTAEG 273
QY 517 QLLKYYKKEKDNALCSIIQNSYADLADIIGSDIIKDYKQKMEENLNKVNKDKKNEE 576
Db 274 NLLLRKNNIYDKLNDLRWSFEDFGDLIMGTDMESIGYSKVKVEKNLRI-----FTGQ 329
QY 577 SLKIFREKWNDEKENVVKMSAVLNK-----ETCKDYDKFKQIPQFLRWPKEWDDF 630
Db 330 NAQLHRKQWNEYKEDIWREWMSSVKKLKGNFVWICEKDPFVNVPEPQIYRWIREWGRDY 389
QY 631 CEKREKIIYSPESFKVECKK-----KQDENTCKNCKSEYKWKIDLKSEYEQVDKY 683
Db 390 ISELPKEL---QKLKEKCDRKFFYTDIKVCTVLPCKNACILYDQWITRKKQKQWVLNKF 446
QY 684 TKQNKMYDNDIDEVKNKEANVYLKESKECKDNFDDKIFNESPNEYEDMC----- 735
Db 447 ---KSVKKVQNTETADIVTAYDILKQELDGFNEVAFENEI-NORDNAYIDLCTVTKNT 502
QY 736 -KKCDIKEYLNEYKPKTKHDIYDIDTFSD--TFGDTPTISINANINEQSGKDTSNIGN 792
Db 503 QKVVTNVENAESAEPISKAKIQAVDSSEKKEKQVGSAGHGNINSAGHNSITGKAVTNGQ 562
QY 793 S-----ETSDSPVSHPEPSDAAINVE-----KLSGDES 820
Db 563 NGNQTPKSNVQRSDIPESASAKNVDPQKYVSERRDDTTTSVTLSIAEAGKNEGLTNGRPS 622
QY 821 SSET-----RGILDINDPSVTNNVNEVHDA-----SN 847
Db 623 ESTVEAYSAGDGTVNSAYIPVKNSENSLVTHKGLGPKSKDNSDNGSTESKESWAPDSN 682
QY 848 TQGSV-----SNTSDITN---GHSESSLNRTTNAQDIKIGRSGNEQSDNQ 889
Db 683 SKGETEMRQDNGKAKATKDSNTSDNTSATSGATGTVDNRINK---SVPEDGNKIYVSK 739
QY 890 ENSSHSSDNGSGSLITIGQVPSDNTQNTYDQNPHRDTPNALASLPDSDKINETEGPDSSR 949
```

Db 740 ENEEDSVNKGATVVGONTNDRNTENDTENN-----LPAPDN-KQSEGATPLS 787  
Qy 950 DSE-----NGRGDTTSN--THDVRTNIVSRRVNSHDPFIRNGMAN-----NNAHH 993  
Db 788 KTELELNERVHRTNTDTLSLKNKGSGEKDLQKHDPFNNDMLNABPNYQOTTDAEGHH 847  
Qy 994 QYITQIENNGIIRGOESAGSNVYKONPKRSNFSSE---NDHKONTQYNSRDTKRVRE 1050  
Db 848 R--DSIQND---KGWRKMKNGAFTKPNPNHNLNHNHNLNNGKLDIKKEYKYRDVNATRE 902  
Qy 1051 EILIKLSONKNNEYSNEYCTYSDESNSSPGPCSRERKKLCCOISDYCLKYFNFYIEY 1110  
Db 903 KLIYSEVRKKNNSILNYSNCSVEDKISS-NTCSREKSNLCCSISDFCLNYPVYSYIEY 961  
Qy 1111 YNCIKSBIKPEYKCFKSEGOSSIPYPAAGGILVIVILLISSASRMGKSNEEYDIGESNI 1170  
Db 962 HNCMKKEFDPSYKCF-AKGSFRKAYPAAGGALLILLILLITSRN-MINKNSEATFNE--- 1016  
Qy 1171 EATPEEN 1177  
Db 1017 ---FEEH 1020  
  
RESULT 3  
T30848  
Duffy receptor - Plasmodium vivax  
C;Species: Plasmodium vivax  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T30848  
R;Fang, X.D.; Kaslow, D.C.; Adams, J.H.; Miller, L.H.  
Mol. Biochem. Parasitol. 44, 125-132, 1991  
A;Title: Cloning of the Plasmodium vivax Duffy receptor.  
A;Reference number: Z20901; MUID:91187056; PMID:1849231  
A;Accession: T30848  
A;Status: preliminary; translated from GB/EMBL/DBDJB  
A;Molecule type: DNA  
A;Residues: 1-1070 <PAN>  
A;Cross-references: UNIPROT:P22290; UNIPARC:UPI0000132B95; EMBL:M61095; NID:g160275; PID  
C;Genetics:  
A;Introns: 19/3; 1006/1; 1032/2; 1057/1  
A;Note: PVDR  
  
Query Match 11.8%; Score 765.5; DB 2; Length 1070;  
Best Local Similarity 25.0%; Pred. No. 1.8e-25;  
Matches 283; Conservative 166; Mismatches 400; Indels 285; Gaps 52;  
  
Qy 272 VEKLNNTKWEKKNKANLWNHMLVNHKG-NISK--ECAIIPAREPOINLWIKENRFL 328  
Db 17 LHKYNNVLLERTITLLECKNEYVKGNGYKLAHGHC---VEEDNLERWLQGTNER-- 70  
  
Qy 329 MEKKRLFLNKKDKVENKYEACPGGCLPCS-SYTSFMKSKTQMEVL-----TNLY 380  
Db 71 -----RSEENIKYK--YGVTELKIKYAQMNGKRSRILKESIYGAHNFGNSY 116  
Qy 381 KKGNGVDKKNFLNDLPKQKNKNDLDDFFKNEKEYDLDLCRY-----TATIIKS 430  
Db 117 MEKGDGDKTG-----BEKGEHKTDKTDNGKANNLVMLDYETSSNGQPGATLDNVLE 171  
Qy 431 FLNGPARK-----DVD-----IASQINVDLGRF-----GCNYSK-SNNKSW 466  
Db 172 FVTGHEGNSRKNSSNGNPDYDIDHKTTISSAI-IN--HAFQNTVMKNCNKKRRRBRDW 228  
Qy 467 NCTGTFNKPFGTCPEPRQTLCLGRYTLHRGHEDY-----KEHLGASTIYEA--- 516  
Db 229 DC-----NTKQDVCPDRRYQLCMKELTNLVNNTDTNPHRDITPRKLYLKRKLYDAVE 283  
Qy 517 -----QLLKYKYEKEDNALCSIIQNSYADLADIIGKSDIKDYKGGKMEENLKV-NKD 570  
Db 284 GDLLLLKLNRY-NKD---FCKDIRSLGDFGDIIMGTDMEGIGYSKVVNENLRSIFGTD 339  
Qy 571 KGRNESLKIIFREKWDENKENVKMSAVLNK-----ETCKDYDKFKIIPQFLRWFK 624  
Db 340 EKAQOR-----RKQWNESKAQIWTAMWYSVKKLGKNTWICKLVNVAVNIETPQYRWIR 394

Qy 625 EWGDDFCERKEKIIYSPESPKVBC-----KKKDCDENTCKNKCSEYKWKWIDLKSEYE 677  
Db 395 EWGRDYVSELPTEV---QKLKERCDGKINTYDKKVKVPCQACKYSDOWITRKQWMD 451  
Qy 678 KQVDKYTKDKKKMYNDIDVKNKEANV-----YLKEKSECKDVNPDFFKIFNESPNYE 732  
Db 452 VLSNKFISVKNAB-----KVQTAGIVTPYDILKQELDFNEVAFENEI-NKRGDGAYI 502  
Qy 733 DMCKKCDBIKYLNEIKYPTKHDYDIDTFSDFG-DGTPI-----SINA 776  
Db 503 ELC-----VCSVBEAK-KNTQEVVTVNDNAKSOATNSNPISQPVDSKAEKVPDSTHG 556  
Qy 777 NINEQSGKDTNTGNSSETSD-----SPVSHE-----PESDAAINVE-----KLSGD 818  
Db 557 NVN---SGQDSITGRAVTGQNGQCTPAESDVQSRSDIAESVSANKVDQKSVSKSDD 613  
Qy 819 -----ESSSTRGILINDPSVTNNVNEVHDSANTQGSV-----852  
Db 614 TASVTGIAEAGKENLGASNSRPSSESTVEANS PGDDTVNSASIPVVSNGENPLVTPYNGLRH 673  
Qy 853 SNTSDITNGHSRSLNRTTNAQDIKIRSG-----NEQSDNQENSSHSSDNGSLT-----903  
Db 674 SKDNSDSDGPAESMANPDSNSK---GETGQDNDMAKATKDSNSSSDGTSSATGDTTD 729  
Qy 904 -----IGQVPSSEDNTQNTYDSQNPHRDTPNALASPSDDKINEIEGFDSSRD-----950  
Db 730 ADVREINKGVPEDRDKTVSGDGGEDNSANKDAATVVGEDRRENSAGSSTNDRKNDT 789  
Qy 951 SENG-----RGDTTSN--THDVRTNIVSRRVNSHDPFIR 983  
Db 790 EKNGASTPDSKQSEDATALSKTESLESTESGDRTTNDTNSLEKNGGKQDKQKDFKS 849  
Qy 984 NGMANNNAHQYIT-----QIENNGIIRGOESAGSNVYKONPKRSNFSEN---D 1032  
Db 850 NDTNPEEPSNDQTTDAEGHRRDSIKND---KAERRKMKNKDFTTKNTNSHHLNNSNLSN 906  
Qy 1033 HKKNIQBYNSRDKRVREIILKLSKQKNNEYSMEYCTYSDESNSSPGPCSRERKCLC 1092  
Db 907 GKLDIKKEYKYRDVKATREDIILMSSVVKCNNTISLEYCNSVEDKISS-NTCSREKSNLC 965  
Qy 1093 QGISDYLKYFNFYISYIYCNICIKSEIKSEPKCFKSEGOSSIPYFAAGGILWIVLSS 1152  
Db 966 CSISDFCLNYPDYVSYEYLSCKMKPEPDSYKCFKTKGFKDKTYFAAAGALLILLIAS 1025  
Qy 1153 ASRWGKSNEYDYGESNIEATFEE-NNYLNKLSRI-----FNQEVQETNISDYS 1200  
Db 1026 -RKMKNDSE-----EATFNEFEYCDNIHRIPLMPNPNIEHMQPSTPLDYS 1070  
  
RESULT 4  
T28652  
erythrocyte binding protein - Plasmodium knowlesi  
C;Species: Plasmodium knowlesi  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T28652  
R;Adams, J.H.; Sim, B.K.; Dolan, S.A.; Fang, X.; Kaslow, D.C.; Miller, L.H.  
Proc. Natl. Acad. Sci. U.S.A. 89, 7085-7089, 1992  
A;Title: A family of erythrocyte binding proteins of malaria parasites.  
A;Reference number: Z20495; MUID:92357776; PMID:1496004  
A;Accession: T28652  
A;Status: preliminary; translated from GB/EMBL/DBDJB  
A;Molecule type: DNA  
A;Residues: 1-1153 <ADA>  
A;Cross-references: UNIPROT:P50493; UNIPARC:UPI0000132B93; EMBL:M90694; NID:g160290; PID  
C;Genetics:  
A;Introns: 20/3; 1087/3; 1114/2  
  
Query Match 11.2%; Score 723.5; DB 2; Length 1153;  
Best Local Similarity 23.6%; Pred. No. 1.2e-23;  
Matches 279; Conservative 171; Mismatches 390; Indels 341; Gaps 51;  
  
Qy 253 DRVKGYNITKCF-SDYYKKNVKNLNNIKKEWKEKKNKANLWNHMLVNHKGNISKECALIPA 311

Db 27 ERWKGIILLSCENEYVKNNGYKL-----ATGHHYM----- 57  
Qy 312 EEPQINLIWIKEMWENFLMEKKRFLNLIKDKC-VB--NKKTEACFGGCRLLPCSSYTFPMKK 368  
Db 58 DNDQIERWLQTDRS---RRVKIEENVKYKYNVEELNTKYEQTKG-----KRINRLIKE 108  
Qy 369 SKTQMEVLNLYKKKSGVGNKPNFLNDL-----PKGNKNDLDDFFKVKBEYD-DLDCRY 423  
Db 109 S-----TYEAQN--VADNNYIDDKANGVEYKTDNKTNGEGARNVWMLDYDISGSH 157  
Qy 424 TATIIK-----SEFLNGPAKNDVDIASQINVNDLRGFG-----CNY 458  
Db 158 PDGLINDVNLGTDEGEFLNLSKGG-DHPYRNKRKRCMCGVINQTFLOKNVMRRCNN 216  
Qy 459 K-SNNEKSMNCTGTFTMKFPGTCEPPRRQTLCLGR-TYLLH-----RGHEE-----DYKE 506  
Db 217 KRKGTRDWDCP---TKK--DVCIPDRRYQLCMKELTNLYNNTKTHSHNDITFLKLNLIKE 271  
Qy 507 HLLGASIYEAQLLYKKEKENALCSIIQNSVADLADIIGSDIHKDYKCKMEENLNK 566  
Db 272 KLTVDAAVEGDLILLKNTNNVYSEDLCKDIKWSLEDFGDIIMGTDMEGIGYSQVVENNLR 331  
Qy 567 VNKKCKNEESLKIIFREKTDENKENVWYMSAVLNK-----ETCKDYDKQKIPOFL 620  
Db 332 VFGTGTQTKQ-----LDRKKWNESKKYIWEATILSVKKKLNGYSAWNCKEDVQINVEPIY 387  
Qy 621 RWFKEWDDDFCE---KKKEKIYSPESPKEBCKKK-----DCDENTCKNCKSEYKKWID 670  
Db 388 RWIREGMDYMSLPKEQRKI-----KEKDRKLYTNLRICTMSPCNDSCLYQWIT 441  
Qy 671 LKSEYKQVDKYTKDKQKMYDNIDVRN-KRANVYLKEKSKCKDVNFDKIFNESPN 729  
Db 442 RKQKQWDLSTKFSVVK-----GQIETENITAYDILKQELNGFNEVMEFENEI-NKRDN 496  
Qy 730 EYEDMC-----KKCDB-IKYLNEIKYPKTKHDIYDIDTFSDTGDTGPIGI-NANIN 779  
Db 497 VIIDICLAADENKNTQEHLLKLLKSPKLETRSHSTIQPMSSSGBAEKVQGLAHGN 556  
Qy 780 E-----QOSGKDTNSETSDSPV--SHPESDAAINVEKLS-GBSSSETGILD 829  
Db 557 DAAVKSITDEAARGDQNGNQTVAESNIKGTDNIEENAAKGVDTYKFTVERSADTRGATD 616  
Qy 830 IN-----DPSVTNN- 838  
Db 617 ITETGEKLTNSYSGSEITVKENIPGDGIVKDVSAAVENSENPLETKHKIPEFSKNSD 676  
Qy 839 -----VNEVHDASNTQGSVNTS---DITNGHSSSLNRTTNAQDIKI 878  
Db 677 NSENGSMFKATSSNPITEAVSSSAEGQVQDSAHRSVNTGRDNSTISAATSDDLGSS 736  
Qy 879 GRSGNEQSDNQENSSHSD-----NSGSLITIG-----QVPSEDN--TQNTYDSQN- 921  
Db 737 GDKRVESLTSIENADDGDPVQGSWNLNLDPSVGAGGKSHIKTEENESQAEIDGNV 796  
Qy 922 -----PFR-----DTPNALASLP-SDDKINE-----IEGPDSSRDSNGRGDT 958  
Db 797 DIAEQRTATITEVQPERPDLSDTDN--GNVPRSGNKQNEGATALSQAESLESNEVHKTI 854  
Qy 959 TSNTHDVRRTNIVSERVRVSHDFTRNGMANNA----- 991  
Db 855 DNTTHGLENKNGNEKDFQKHDFMNDMLNDQTSDDQTSDDQTSDDQTSDDQTSDDQTS 914  
Qy 992 -----HHOYITQIENNGIIRGOESAGNSVNY 1018  
Db 915 DQTSDDQISSDQTSDDQTSDDQTSDDQTSDDQTSDDQTSDDQTSDDQTSDDQTS 972  
Qy 1019 KDNPKRNFSENDHKQNIQEVNSRDTKRVREIILKSKQKCNKNEYSMEYCTYSDBRNS 1078  
Db 973 SNGNELYSHNNLNKLNIDQYEHRDVKATREKILMSEVKNCKNRASLKYCNTIEDRML 1032  
Qy 1079 SPGPCREERKKLCCQISDYCLKYFNFYSTIYXNCIKSEIKSPYKCFKSEGOSS---IP 1135

Db 1033 S-STCSRERSKNLCCSISDFCLNLYFELYFYFVNCMKKEFEDSSYECF-TKGSSTGIGIV 1090  
Qy 1136 YFAAGAILVIVILLASSASRMGKSNBEEYDIGESNIEATPEE 1176  
Db 1091 YFATGGAFILILLFFSKKNVANDYEE-----EATFDE 1123  
RESULT 5  
A35970  
erythrocyte-binding protein - Plasmodium knowlesi  
C:Species: Plasmodium knowlesi  
C:Date: 14-Dec-1990 #sequence\_revision 14-Dec-1990 #text\_change 09-Jul-2004  
C:Accession: A35970  
R:Adams, J.H.; Hudson, D.E.; Torii, M.; Ward, G.E.; Wellem, T.B.; Aikawa, M.; Miller, I.  
Cell 63, 141-153, 1990  
A:Title: The Duffy receptor family of plasmodium knowlesi is located within the micronem  
A:Reference number: A35970; MUID:91004213; PMID:2170017  
A:Accession: A35970  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-778 <ADA>  
A:Cross-references: UNIPROT:P22545; UNIPARC:UPI000016BFA8; GB:M68518; GB:M37513; NID:916  
C:Keywords: transmembrane protein  
Query Match 10.1%; Score 655; DB 2; Length 778;  
Best Local Similarity 24.2%; Pred. No. 6.4e-21;  
Matches 192; Conservative 131; Mismatches 276; Indels 194; Gaps 26;  
Qy 532 CSIQNSYADLADIIGSDIHKDYKCKMEENLNK-VNKKCKNEESLKIIFREKWDENK 590  
Db 2 CKDIRGLGDFGDIIMGTNMEGIGYSQVVENNLRQVFGTDEKAKQD-----RKQWNESK 56  
Qy 591 ENVWKMVSAVLNK--KE-----TKQYDKFQKIPQFLRWFKEWDDDFC-----EK 633  
Db 57 EHLWRAMFESISRLKEKFKFWICKKQDVTLKVPQIYRWIREWGRDYMSKLPKEQGLNEK 116  
Qy 634 RKEKYSFSPFVKECKDQDENTCKNCKSEYKQKWLIDKKSEYKQVDKYT--KDNKKM 691  
Db 117 CASKLY-YNNMAI-----CMLPLCHDACKSYQDQWITRKKQWDLSTKFPSSVKTKQIG 169  
Qy 692 YNIDBEVKNKEANVY--LKEKSKCKDVNFDKIFNESNEYEDMCKCKDEIKYLNSEIKY 749  
Db 170 TENI-----ATAVDILKQELNGFKBATFENEI-NKRDNLNHLCPVVEEAKRN---- 217  
Qy 750 PKTKHDIYDI-----DTFSDTTFDGG--TPISINANINEQSGKDTNSTGN 792  
Db 218 --TQENVKNVSGVESKAASSNPITEAVKSSSGEGKVQEDSAHKS VNKGEKGSSTNEADP 275  
Qy 793 SETSDSPVSHPESDAAINV-----EKLSDGESSESTRGILDIND----- 832  
Db 276 GSQSGAPASRSVDEKAGVPALSGAGHDKVPPAAEAAATBSAVLHSAADKTPNTVTENKEG 335  
Qy 833 -----PSVTNNVNEV-----HDASNTQGSVNTSDITNGHSES 865  
Db 336 TQWDGAAGDGGKAPGPTVSSDVPFVGKSGSGPSTASHALAGENGVEHNGTDT---EPKE 392  
Qy 866 SLNRTTNAQDIKIGRSGNEQSDNQENSSHSSNSGSLITIGQVPSEDNTQNTYDSQNPHRD 925  
Db 393 DGEKADPQKDIEV--KGKQDTRDSQSGSLGPHPTDERATLGETHMEKDTETAGSTL--- 446  
Qy 926 TPNALASLPD-----DKINE-----IEGPDSSRDSNGRGDTTNSHTVARTIVSE 973  
Db 447 TPEQNVSVASDNGNVPSGKNQNEGATALSQAESLKSNESVHKTIIDNTTHGLENKNGNE 506  
Qy 974 RVVNSHDFTRNGMANNA----- 991  
Db 507 KDFQKHDFMNDMLNDQASSDHTSSDQTSDDHTSSDHTSSDHTSSDHTSSDHTSS 566  
Qy 992 -----HHOYITQIENNGIIRGOESAGNSVNYKDNPKRNFSENDHKQNIQEVNSR 1043  
Db 567 DQITIDTEGHR--DNVRNPEIKSSDMSKGD FMRNSNSNELYSHNNLNKLRNDRDQYEHR 624  
Qy 1044 DTKRVREIILKSKQKCNKNEYSMEYCTYSDBRNSPGPCREERKKLCCQISDYCLKYF 1103





A;Residues: 1-2924 <BAR>  
A;Cross-references: UNIPROT:Q25733; UNIPARC:UPI0000079077; EMBL:U27338; NID:G914918; PID:G914918  
C;Genetics:  
A;Gene: EMPL  
A;Introns: 2476/3

Query Match 8.6%; Score 555.5; DB 2; Length 2924;  
Best Local Similarity 22.7%; Pred. No. 5e-16;  
Matches 280; Conservative 182; Mismatches 434; Indels 337; Gaps 65;

Qy	48	ELANNHNTIYDSYEDVNNKLIINSFVENKSVKKGRSLSFNNKTKSYDIIPPSYSYRND	107
Db	1674	KINNELKGN-----GKDFNGKCN-----VKKNGAVI-----	1706
Qy	108	KFNSLSNEDNSGNTNNSNFANTSEISIGKDNKOYTF-IOKRTHLPACGIKRKSIIKWICR	166
Db	1707	KFGQTYNSVN-----NINKC-----KDNQNERFKIQKNFKYIGIRKDL-----	1749
Qy	167	ENSEKITVCVPPDRKIQLCVANF-----LNSRLTMEKFKBEPLISVNTAEKLLYNKNE	219
Db	1750	-----CIPPREHEHCLDLSMLGRTTISDSSALLKKIQEAAKSERDDIIRKLLQNS	1801
Qy	220	GKDPSPICNELRNSFSRFSFGD-----DMDFGGNTDRVKGY---INTKPSDYK-----	268
Db	1802	CDEHRI--CDAMKYSFAD-----LGDITRGRDLWNKNSKQKGLQKRLEYAFINIYNKLQND	1855
Qy	269	-----EKNVKELNIIKKEWKEKNKANLWNHVLN-----HKG-----	300
Db	1856	KNKEKDRPKYQLRSOWDANKRHIIWNMTCPNAPDDAKPLKGNPNDTSGSSSKGIMTT	1915
Qy	301	-----NISKECAIIPABE---PQNLMTKEWNEF---LMEKRLFLMIKOCVEN-----	345
Db	1916	HSNCGYDKE---PPDYIPIQPFRWQWSESPCKLLNEEMEQUEFKTCECKKNSITCE	1971
Qy	346	--KYEACFGGRLPCSSYTSFMKKSK-----TQNEVLTNLYKKNSGVDKGNFLNDLPKK	399
Db	1972	DDRGNTNC--ENCRNQCEKYKGLIHNWKLGDFKYKBYIYNEIYNNKNSKINSNEYFKPELEK	2030
Qy	400	-----NNKNDLDDFFKN-EKEYDDLCDC-----	421
Db	2031	LKDCKELNSDDKCIDBEATHCTKYKFSNENKHNNTAFKNPPKBEKACKCDAPDPLDN	2090
Qy	422	--RYTATIISFLNGPAKNDVDIASQINVNDLGRFGCNYK--SNNEKSNMCTGTFTTNKFP	477
Db	2091	CPKDSATYEK-----ACNTL-LPTKL-----CESTFNDDDSWDT--SFVQTSP	2132
Qy	478	-----GTCEPPRRPTCLGRYVLLHRGHE---EDYKEHLLGASIVYEAQLLKYYKEKBNAL	531
Db	2133	RDNTGVLVPPRRQICLKNITTKLRSITEKIDDFKAEIIMTSAYNEGKLLCELYK-KORDVT	2191
Qy	532	CSIIQNSYADLADIIGSDIIKDYGKKEEENLNKVNKKRNEESLKIIFREKWWDNKE	591
Db	2192	LQAMKYSFYDYGDIVKGTDLISTAPLDKLTKNLVLLKGDTNE--IKEDRGKWWTSNRT	2249
Qy	592	NVWKVMSAVLK-----NKETCKDYDFKIPQPLRFWKEWGDGFCERKEKYSFESFK	645
Db	2250	RVWHAMLCGYAKGKLEERDCSLPD--DNTHQFLWFRWSEHFCAQRQK---LNFVVK	2304
Qy	646	VECKKQDC-----DBNTCKNCSYKKWIDLKSEY-----EKQVD-----	681
Db	2305	RECAQAQIIIEYGTIDPPVCBEAQYRDIITRKIQEYRLNLYQYNTFNFEKAEVTKAP	2364
Qy	682	KYTKDKN-----KKQYDNID--EVNKEANVYLKESKECKDVNDDK	722
Db	2365	EYFNDKNDKNCLSKYDIEKKWKMYDSFDONDKNKICIQRIKPK-RPPKKVKPEEE	2423
Qy	723	IFNESPNYE---DMWCKCDEIKYLBIEKPKTKHDIYDIDTSDTFCGDTGPIINANIN	779
Db	2424	---HTPSEQDTPPPLPKPDOL-----PPPAEPPFNRLLEKTIPIFGIALALGSIAP	2472
Qy	780	EQSGKDTNNGNSETSDFSVSHEPESDAAINVEKLSGDE-----SSSETRGILDINDPSV	835
Db	2473	LFLKCKTKSSVGNL-----FOILHPIKSDYDIPT-KLSPNRYIPVTSKQYKQ-----	2518

Qy 388 DKNFLN-----DLFKNNKNDLD-----DFFKNEKEYD-----DLCDCRY 423  
Db 1914 ESCQTOHNPNGTKPCDEIPPHSDTELDEQDTTTTDDNSDKIYDTKPPCPKQVEDTKE 1973  
Qy 424 T-----ATIKSFLNGPAKNDVDIASQINVDLGRFGCNVY-----SNNEKSWN 467  
Db 1974 TEXPKVLPGPDPDACEIVEILNGO-----DGTKKIB-----ECNTKYPTKNDYFGWN 2021  
Qy 468 CTGTTNFKPGTCEPPRRQTLCL-GRTYLLHRGHEEDYKEHLGLGASIEAQLLYKYKE- 525  
Db 2022 CTDKVINREEGSCMPRRQKLCIHNLHLSEKATETELRAKAFIECAAETFLWLWDYKED 2081  
Qy 526 -KDNAL-----CSIIQN-----SYADLADIKSGDIIDKY-----555  
Db 2082 KDEKTEGGISDDPDPPOKKEGGTIPEDFKRQMFYTYG DYRDLFGDTDISKGHGKES 2141  
Qy 556 -YGKQMEENLKNVKKRNEESLKIIPREKWDENKENVKNVMSAVLK-----NKETCK 608  
Db 2142 ALGKKI-DSLFK-NGDQKSPSGKTP--TEWWDYGDPIKNGWVCGLSHHIKNGKEQLR 2196  
Qy 609 ----DVKFOKI-----PQFLRWFKWGGDDFCEKKEKIKIYSPESPKVECKKDC- 653  
Db 2197 KNLTDNNKYTKISKLEDFASRPQFLRWFIWEGDQFC---RERVVKINQLKTKGNEVECG 2253  
Qy 654 -DEN-----TCNKCSEYKWKIDLKSEYKQVDKYTKOKKKMYD-----NIDEVKNKEAN 704  
Db 2254 SQENGKKEACKNACEAYKSLWKQWQYEQOTAKFKDKKDKFGDGTSAEVDVAAVSSVH 2313  
Qy 705 VYLKESK-----ECKQVNPDDKIFNESPNEBYEMCK-----736  
Db 2314 EYLQEEELKNLCTKGDCACMEKPSAQDETELGLGNYPPEAMDPKPEIGERCKCALPSEP 2373  
Qy 737 -KDE--IKYLN-----PKTKHDIYIDTFSDTFGDGTPIISINANIN 779  
Db 2374 MSCVEQIAKHLREKAENKVIYESSLGTGTPAKSKNDCTKID--EAIKGDNGSKIINKSI- 2430  
Qy 780 EQSGKDTNVTGSETSDSPVSEHPESDAAINVEKLSGDESS---SETRGILINDPSVT 836  
Db 2431 ----LDSTFASNEQSEKDATRLKIGKQWQFNKNGTGTGLYVPPRRKDMCFND---L 2482  
Qy 837 NNV--NEVHDASNTQGSVNTS---DI--TNGHSESLNRTTNAO-----DIKI 878  
Db 2483 KNTQFNEVQDSNLSLEKIQHVAKNEGIDILKLNPDQNAFSELCAMKYSFADLGDIIIR 2542  
Qy 879 GRSGNEQSDN-----QENSSHSDNSGSLTIQGVPESDNTQNTYDSQNPHRDTPNALA 931  
Db 2543 GRSKIDPTNNKIEKELQKIFKQIODDNASLSKMELP--ELREKWDAN--RKEVWNAMT 2598  
Qy 932 SL-PSDDKINEIEGFDSRSRSENGRDTT---SNTHDVRTNIVSERRVNSHDFIRNGM 986  
Db 2599 CVAPNDABL-----KKKQNPNGKSOIILASQTEQTKC-----SHD-----2634  
Qy 987 ANNAHQYITQIENNGIIRGQESAGNSVNYKDNPKRSNFSSENHDKKNIQEYNSRDTK 1046  
Db 2635 -SEPPDYDI-----PERYF-----LQWSEYCYCK 2659  
Qy 1047 RVREIILKSKQKNKNNEYCYTSDERNSSPGCSREERKKLCCQISDYCLKYPNF- 1105  
Db 2660 ALKE-----KNDEMKNDCSK--CI-----KSGATCEKEEDKECKECNCKKEYKNIV 2705  
Qy 1106 --YSIEYNNICKSEIKSPKCPKSEGOSSIPYFAAGILVIVILLSSASRMCKSNEEY 1163  
Db 2706 DRWQSEF-----DQOQLYKGLYQDTRTHGP-----STARRNPSTIEFT 2743  
Qy 1164 DIGESNTEATFEENNYLNKLSRIFNOBVOETNISDYSEYNY 1204  
Db 2744 QKLEDSQNDYSAKYLIDISTHCTDYKSET---DSRESNY 2781

RESULT 10  
T28432  
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)  
N;Alternate names: erythrocyte membrane binding protein 1 (EMP1)

C;Species: Plasmodium falciparum  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T28432  
R;Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herrfeldt, J.A.; Peterson, D.S.  
Cell 82, 89-100, 1995  
A;Title: The large diverse gene family var encodes proteins involved in cytoadherence an  
A;Reference number: Z20487; MUID:95330813; PMID:7606788  
A;Accession: T28432  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-3078 <SUX>  
A;Cross-references: UNIPROT:Q26031; UNIPARC:UPI00000820C2; EMBL:L40608; NID:9886374; PID  
C;Genetics:  
A;Gene: var-1  
A;Introns: 2611/3  
  
Query Match 7.7%; Score 498.5; DB 2; Length 3078;  
Best Local Similarity 18.7%; Pred. No. 1.4e-13;  
Matches 317; Conservative 180; Mismatches 520; Indels 661; Gaps 67;  
  
Qy 90 NKTYSYDIIPPSYSYRNDKFNLSNEDNSGNTNSNFPANTSE-ISIGKDNKOYTFIQKR 148  
Db 1309 NKEYTEKQPPPEYATACDCINRSQTEEPKKKEENVESACKIVKILEGKNG-----R 1360  
Qy 149 THLPACGIRKRSIKWICRENSE--KITVCPDRKIQLCVANFLN-----SRLETMEKEKEI 202  
Db 1361 TTVGECNPKESYPDWDCNNIDISHDGACMPRRQKLCLYYTAHESQTEINIKTDDNLKDA 1420  
Qy 203 FLISV-----NTEAKLLYNKNEKQPSIFCNELRNFSFSPRSSFIGDDMD 247  
Db 1421 FIKTAAAEFLSWQYKYKNDSEAKIL---DRGLIFSQFLSRMMYTFG DYR-----D 1469  
Qy 248 FGNTDRVKGYINTKPSDYKXKNVEKLNLI-----KKWEKKNKANLNHNM 294  
Db 1470 ICLNTD-----ISKQNDVAKAK--DKIGKFFSKGSKSPSGLSRQEWKWTNGPEIKWGM 1522  
Qy 295 IV-----NHNKGNISKECALIPABEPQINLWIKENWENFL 328  
Db 1523 LCALTKYVTDTONKRIKNDYSYDKVQNSQNGNPSLEEF---AAKPQFLRMWIEWGEFC 1579  
Qy 329 MEKRLFLNLIKQCVENKKEVEACFGG---CRLPCCSYTSFMKSKTKQMEVLTN-----378  
Db 1580 AEKQKKNIIKACNENINSTQCNDAKHRCNQACRAYQEVENKKEKFFSGQTNFVLKAN 1639  
Qy 379 -----LYKKNSGVKNNFLNDFKKNKNDLD-----DFFKNEKEYD 416  
Db 1640 VQODPEYKGYEKDGVQPIQGNAYLLQKCDNNKSCMDGNVLSVSPKEKPFKYAHKYP 1699  
Qy 417 DLCDCHYTATIISKFLNGPAKND-----VDIASQINV--NDLRGP-----GCYKSN 461  
Db 1700 EKCDC-YQGHKVPISIPPPPPVQVQPEAPTVTVDVCSIVKTLFKDTNNFSDACGLKYGKT 1758  
Qy 462 NEKSNMCTGTFTNKFPGT-----CEPPRQTLCLGR-----492  
Db 1759 APSSWKCIFSDYKSGAGATTGKSGDSGSICIPRRRLYVGLQEWATALPOGEAAPS 1818  
Qy 493 -----TYLLHRGHEEDYKEHLGLGASIEAQLLYKYKEKDNALCSCI 534  
Db 1819 HSRADDLRNAFIQSAAIETFFLWDRYKEBKQPQDGSQQALSQLS-TYSDDEEDPDKL 1877  
Qy 535 IQN-----SYADLADIKSGDIIKDYKKEENLKNVKKRNEESLKIF 581  
Db 1878 LQNGKTPDPLRMFLMYTLGDYRDLVHGNGTSD-SGNTNGSNNNNIVLEASGNKEDMQKI 1936  
Qy 582 RE-----KWDENKENVKNVMSAVL-----KNKETC-----607  
Db 1937 QEKIEQILPKNGGTPLVPKSSAQTPDKWNEHAESIKGMI CALTYTEKNPDTISARGDEN 1996  
Qy 608 ---KD---YDKF-----QKI-----616  
Db 1997 KIEKQDEVKEKFGSTADKHGTASTPTGTGYKTDYKVKLEDTSGAKTPSSASDTPLLS 2056  
Qy 617 ----PQFLRWFKWGGDDFCEKKEKIKIYSPESPKVECKKK-----651

Db 2057 DFVLRPPYPRYLEWQNFCKRKHKL---AQIKHECKVBENGSGRRGGITROYSGDGE 2113  
QY 652 -----DCBENTCNKCSYKKWIDLKSEYKQVDKYTKDKNKKMYDNIDE 697  
Db 2114 ACNEMLPKNDGTVPDLKSPCKPCSSYRWKIESKGFEKQKAYPQQDKCVNGS--- 2170  
QY 698 VPKNEANVYLK-----BKSKE-----CKDVNF-----DDKIFNESPN-----EYED 733  
Db 2171 --NKHDNGFCETLTJTSKAKDFLTKLGPCKPNNVEGKTIIPDDDKTFKHTKDCDCLAFSV 2228  
QY 734 MCKK--CDEIKYLNEIKYPTKHDIYDI-----DTPSDTGDGTPPISINA 776  
Db 2229 NCKOBCDNGKGT--CRNKNISDATDIENGVDSTVLEMRVSADSKSGFGDGLNACRG 2286  
QY 777 -----NINEQSQKD----- 786  
Db 2287 AGIFGIRKDBWKCRNVCGYVCKPENVGAEKGHIIQIRALVKRWVYFFPDYNNKIKH 2346  
QY 787 --TSNTGNSETS-----DSPVSHPEPESDAAINVEKLSGDESSSETRGILD--- 829  
Db 2347 KISHRIKNGEISPCIKNCVKWVDQKKEWKIEITERPKDQYKNDNSDDNVRSFLETLIP 2406  
QY 830 -INDPSVNNVNEVDASNTQG--SVNSTSDITNGHSESSLNRTNAODIKIGRSGNQSD 887  
Db 2407 QITDANAKNVIKLSKFGNSCGCSASANEQNGEYKDAIDCMLKCLKDKIGEC--EKKH 2464  
QY 888 NQENSSSHSSNSGSLTI-----GQVPSD 911  
Db 2465 HQTSDSCSDTPOPOTLEDETLDDIETEAKGNMMPKICENVLKTAAQDEGCVPAEN 2524  
QY 912 N-----TQNTYDSQNP 923  
Db 2525 SERPAATDSKETPEQTPLVKPEEAVPEPPPPQEKAPAPPOPPPTPTQLLDNPH 2584  
QY 924 -----RDTFNALASL-----PSDD-----KINE----- 941  
Db 2585 VLTALVTSTLAWSGVIGFATFTFYLYKTKTSVGNLFQILQIPKSDYDITPKLSPNRYI 2644  
QY 942 -----IEGFDSSRDS--ENGRGDTTSNTHVRRNIVSERVNSHDPIRNG 985  
Db 2645 PYTSGKRGKRYIYLEG--DSGTSGYTDHYSDITSSSEYEYEMDI-----NDIYVPG 2695  
QY 986 MANNAHQYITOIE-----NNGIIRGQBEA--GN--SVNYKDNKPNFSSND--- 1032  
Db 2696 SP-----KYTLIEVLEPSENNNTASGNNTTASGNNTTASGNNTSDTQNDIQNDGIP 2749  
QY 1033 -HKKNIOEYNSRDTKRVREIILSKQNCNNEYSMEYCTYSDBERNSSPGPCSRE-ERKK 1090  
Db 2750 SSKITDNEMN-----QLKDBFISQYLQSEPNTEPNM--LGYNVDNNTHTPTSHNVBEKP 2802  
QY 1091 LCCOISDYLKYENFYISYIYVNCIKSEIKSPYKCPKSEQSSIPYPAAGGILVIVILL 1150  
Db 2803 FIMSIDHR-----NLFSGEYN-----YDMFNS--GNFPI----- 2830  
QY 1151 SSASRMKSNEEYDIGSNTSEATPENNYLNKLSRIFNQSVQETNISDYSEYNNEN 1208  
Db 2831 ----NISDSTNSMDSLTNSNHSFNDKNDLYSGIDLINDALSGNHIDIYDEMLKRKEN 2884

RESULT 11  
T28157  
erythrocyte membrane protein 1 - malaria parasite (Plasmodium falciparum)  
C;Species: Plasmodium falciparum  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T28157  
R;Yang, S.Q.; Wollish, W.S.; Gut, J.; Wu, J.; Ahn, J.; Petersen, C.; Fujioka, H.; Aikawa  
submitted to the EMBL Data Library, July 1995  
A;Description: The molecular cloning and DNA sequence analysis of Plasmodium falciparum  
A;Reference number: 220479  
A;Accession: T28157  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA

A;Residues: 1-2212 <YAN>  
A;Cross-references: UNIPROT:Q94657; UNIPARC:UPI00000829E8; EMBL:U31083; NID:g1517813; P1  
C;Genetics:  
A;Gene: EMP1

Query Match 7.5%; Score 483; DB 2; Length 2212;  
Best Local Similarity 20.5%; Pred. No. 4.5e-13;  
Matches 310; Conservative 166; Mismatches 450; Indels 584; Gaps 73;

QY 36 DESDLSRVNSPELNNHKTN-----IYDSYED-----VNNKLINSFV 74  
Db 701 DELLDELHEKAKOCIVNHKDNCCPDLLSDSEDEEDIPQONPCAKPSGAHRALVNKVA 760  
QY 75 ENKSVKKKRL--SFNNKTKSYDIIIPPSYSYRNDKFNSENEDESGNTNSNFA---N 129  
Db 761 SNMHKKRQLVNRGVSSKLG-DAAGVYR---KSGTTIKLKDICSITNDSHNAKRGH 815  
QY 130 TSEISIGKONKQVTFQKPTH--LFCAGIKRKSIKIKWCENSEKIVTCVPRKIQLCVAN 187  
Db 816 TDQPCRKDKSKSEMP---RTEDGKPKAGFISKTVK-----DIYMPRRQHGFCTSN 862  
QY 188 FINSRLTMEK-----FKEIFLISVNTA---KLLYNK-NEGKDSIF 226  
Db 863 L--EYLQTTNKLNGNDINGNPNIINDSFLGDVLFAANYEADFIKMYKKONDYKDNATI 920  
QY 227 CNELRNSFSDFRSFSGDDM---DFGN--TDRVKGVIINT-----KPSDYKKNVEKL 275  
Db 921 CRAMKYSFADLGDIIKGTDMWDSGESKTRDKLRIFDIKKHPIKIIYEDT--PY 978  
QY 276 NNTKQWWEKQKLANLNMH--IVNHKNISKCAIIPABE--PQINLWIKENNFLEKK 332  
Db 979 TKLRBDWEANRKKIWEAMQCPENGSPFCKSYHTPLDDYIPQRLRWMTWAEWFCQK 1038  
QY 333 RLFLNIKDKC-----VENKYEACFGCRLPCSSYSYSPMKSKTOMEVLTLYK--- 381  
Db 1039 KOYGLVSACNGCKDEGKVCNKSQC--TSCMOACENYKNFINTWKEQWDMBIKYL 1097  
QY 382 -----KQNSG-----VDKN-----NFLNDLFPKKN----- 401  
Db 1098 LQAQTTAANGPPTYSGLDENKPVVNFLELYKENGKIGNPRDTPRAKRSKRETAPA 1157  
QY 402 ---KNDL-----DDF-----FKN--EXEYDDLCRY 423  
Db 1158 SVAKNDVYSTAAGYVHQEMPHMECKTQTECEKTDQYNNYNTFKNPPPYQKDACICN- 1216  
QY 424 TATIIKSLNGPKND-----VD----- 441  
Db 1217 -----TRPPPKDSRKRSEDSDBEEKYKTKVBEKATEDAVDTGPPAPKEATTLD 1268  
QY 442 ---IASQINVDNLRGFC--NYKSNNEKSNWCTGT--FTNKFPPT----- 479  
Db 1269 VCEIVAGVLTKENLEN-ACPTKYPKAPTWSKCIPTKNTNAATGSESGSGNALQAKRA 1327  
QY 480 -----CEPPRRQTLCLGRTYLLHRGH-----EBDYK 505  
Db 1328 TVESGPTVTSNGSICIPRRRLYIQLHDWASGNTVWSQAQTPGGTSPSGKETPS 1387  
QY 506 EHLHGASIVEAQLKY-----KYY-----EKDENALC----- 532  
Db 1388 DKLRTAFIQSAATETFFLWDRYKKEKIEKKKVKANGGLVPSLANGGPPQPGVTGDSQP 1447  
QY 533 SIION-----SYADLADI IKG--SDIIKDYTG-----KMEENLNKNV 568  
Db 1448 SKLQQTGVIPPPPLRQMFYTLGDYADIFFGKNDIVIDTKNGDKDIAERKKIKDAIERVL 1507  
QY 569 KDKKRNEESLKIPREKWDENKENVKMSAVL-----KN 603  
Db 1508 KNADSQPPSDE--KRQTWEGNGEHIWNGMICALTYSKDEKGTPLKQNEGLKSALWDEKN 1566  
QY 604 KETCKD---YDK-----FKIPQFLRWFKWGDGDFEKKR 635  
Db 1567 KKP-KQKQYDKVKLDENSGTSPKTNHVPPTPLTNFISRPYPFRYLEWGTFCRERK 1625

QY 636 EKIYSPSPKVC-----KK-----KOCDE-----NTCKKCYSEYKWI 669  
Db 1626 KRL-----EKIYVECWDEGKKQKSGDGEDCEEIRKODYSTVRDYCEPCYCRFYKRWI 1682  
QY 670 DLKSEYKQVDTYKDKNKKMVDNIDEVN-----KEANVYLKE--KSKECKDVNFD-- 721  
Db 1683 EKKDEYDKQKEAYNNQKTARNNNDNAFSTLDTCTTAGDFLOTLKNGPCNDNVDD 1742  
QY 722 -----KIPNESN--EYEDMCKCDEIKYLBNIKYPKTKHDIYDITFSDTGGDGTPI 775  
Db 1743 GENKKIPDENGDTFKYQYCGTCS-----LNGFK-----CNGDDCVRVTN 1782  
QY 776 ANNEQO-----SGKDTNNTGNSSETSPVSHPEP-----DAAINVEKLGDRSS-- 821  
Db 1783 VTCNGSNRTTITADDTKNGNSAEINMLVSDOINGNGFNDLEACKNANIIFKGIKENKW 1842  
QY 822 -----SETRGILDIND-----PSVTNNVNEVHDASNTQGSVSNSTSD 857  
Db 1843 KCVYFCKSDVGLKKNNDIDONQIILIRALFKRWLEFLDYDKIRKLPNCINNGEKAI 1902  
QY 858 ITNGHSESSLN--RTTNAQDIKIGRSGNEQSDNOENSSHSDNSGSL--TIQVPSSEDTQN 915  
Db 1903 CTNGCVBQWIMQKRTETWNIK--NRFNEQYGGDDTEMKSSFRSFLVDLIRQIAA----- 1954  
QY 916 TYDSQNPRTDTPNALASLPDSDDKINEIEGDFSSRDSSENGR--GDTSTNTHDVRRTNIVSER 974  
Db 1955 TIDKGN-----HNGLVKLVRKVKCN-----CGNNSQNGKEGE----- 1986  
QY 975 RVNSHDFIRGMANNAAHQVITQIENNGIIRGOESAGNSVNYKONPKESNFSS----- 1029  
Db 1987 -----ENDVLCLLQKLEKAEKCKONPETSIGIPQOPCEV 2021  
QY 1030 -----ENDHKNTQYNSRDTKRVREBIKLSKQNKCNNEYSMEYCTYSDERNSSPQPC- 1083  
Db 2022 SPNHIEDEQPLEBEENTVEHPKICDDVLKTPQ-----PEEPGETC-----BESPGPTD 2071  
QY 1084 ---SREERKK 1090  
Db 2072 VKEEBEEKK 2081

RESULT 12  
T09127  
Probable erythrocyte-binding protein MAEBL - Plasmodium yoelii  
C;Species: Plasmodium yoelii  
C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C;Accession: T09127  
R;Kapke, S.H.I.; Noe, A.R.; Fraser, T.S.; Blair, P.L.; Adams, J.H.  
Proc. Natl. Acad. Sci. U.S.A. 95, 1230-1235, 1998  
A;Title: A family of chimeric erythrocyte binding proteins of malaria parasites.  
A;Reference number: 216577; MUID:98115903; PMID:9448314  
A;Accession: T09127  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1701 <KAP>  
A;Cross-references: UNIPROT:O61164; UNIPARC:UPT000007D433; EMBL:AF031886; NID:G2947227;  
A;Experimental source: subspecies yoelii; strain YM  
C;Genetics:  
A;Gene: maeb1  
A;Introns: 62/1; 1648/1; 1674/2; 1697/1  
C;Keywords: alternative splicing; cell binding; erythrocyte invasion

Query Match 7.0%; Score 456; DB 2; Length 1701;  
Best Local Similarity 19.9%; Pred. No. 4.8e-12;  
Matches 273; Conservative 192; Mismatches 510; Indels 398; Gaps 57;

QY 33 NRODESSDISRV-----NSPELNNHNTNIYDSYEDVNNKL-----INSFVENKSVKX 81  
Db 436 NQNIETPDINRVKIVKRNSELSNLYKNDKFKNSDINKGMAINMMDINEIKENSKLQT 495  
QY 82 KRSLSPFNNTK-----SYDIIPSPYSYVRNDK-----FNSLSENEDNS--GNTNS--NNF 127  
Db 496 NK-----GNETKTKYGLYNYPTIPISYLQIHHQWELKYNMDSNSFTSFHNTNAPTHY 550

QY 128 ANTSEISIGKDNKQYT-----FIQKRTHLFACGI 156  
Db 551 EGNKSFSTGVNNKRENTYGTQDINLNRNNYNNQPKNPQABYMDRFDIEKNHIY----- 606  
QY 157 KRKSIKW-----ICRESEKIITVCPDRKIQLCVAFNLSRLB--TWEKPK 200  
Db 607 ----IDWKQDGKYGSGKLKYNIIISHETADTIOSLLITDKDDICPNHYSFGRAQSCPNYG 662  
QY 201 EIFLI-----SVNTEAKLLYNKNEGKPSIFCNELRNSFSDFRSSFIGDDMDFGNTDRVK 256  
Db 663 KSIIVKTPESINGNEHL--NSN-----FLNEIRT----- 689  
QY 257 GYINTKFSYDYKKNVE-----KLNNIKKWEKWN-----KANLWNH----- 293  
Db 690 GYLNK-----YMKSNVELPEYKSLAMHGDLSVCPKSWDEENLYKKNRDYNNDMCKSTV 744  
QY 294 -----MIVNHK-----GNISKECAIIPABEPOINLWIKENWENFLM 329  
Db 745 MKSTIPLKMPDYKTKKLYGLYGLGRLGNSISKYKNIFKSPNNITL-----PM 795  
QY 330 EKKRLPLNIKD-----KCVENKKYEACFGGCLPCSSYTSFMKCKSKTOMEVL----- 376  
Db 796 FNPSSIKNLLDCLSYCYCLGPLENAYNNKCF-----RSLPAYFNHETNECIIIGTHE 848  
QY 377 ---TNLYKKKNSGVDRNFNLFKN-----NKQDLDDFFKNEKEYDDLCDCRYTATI-- 427  
Db 849 QERNNNCRTRRSDDTKPNCQN--VRKNISTQNTWYVTSFIR--PDYEKCPPRYPLKFKS 904  
QY 428 -----IKSFLNGPAKNVDI-----SFTSLTNDI 1002  
Db 905 FGKIDETCKCKSLIN--KXNIINIPFSSCLEYMFIMYPSVLQRTKKNYWGVMVASES 962  
QY 443 ASQINVNDLURGFCNYSNNKSNWCTGTFTNKPFGTCBPPRRQTLCLGRTYLLHRGHE 502  
Db 963 VNSSNLNNAKG--ECYY--INEXP--NCVIDKNHF----- 1002  
QY 503 DYKHEHLLGASIVYEAQL-----LKYKYEKDENALCISIQ-----NSYA 540  
Db 1003 DFNQINLVKLDLVINNDQSSSHNRKNTYPIENSESTIVRKHNSAPEHFRSLKINSYT 1062  
QY 541 DL-----ADIIGSDI IKDYGKMEENLNKVNKDKGNEBSLKIIFREKW---WDEKENVW 594  
Db 1063 PNRGENFAKESDSTNTDESKMDEVIRK--REBAAKNAKBIIRFEBEAQKAAWAKKAEER 1121  
QY 595 KWSAVLNKNETCKOYDKFKIPQFLRWFKWGGDDFCERK-----EKIYSFESFKVECKK 650  
Db 1122 KKAEEAVKAAEERKRIEAEKKAEE-----EERKRIEAEKAAEERKRIEAEK 1167  
QY 651 KOCDE-----NTCKNKCSEYKWIIDLKSEYK---QVDKYTKDNKK--MYDNIDEVKNK 701  
Db 1168 KAEERKIIEAAKAAEERKRIEAEKAAEERKRIEAEKAAEERKRIEAEKAAEERK 1227  
QY 702 -EANYVLKESKCEKDVN--FDDKIENSPNEVEDMCKCDEIKYLBNIKYPKTKHDIYD 758  
Db 1228 AEAAKAAEERKRIEAEKAAEERKRIEAEKAAEERKRIEAEKAAEERKRIEAEKAAE 1282  
QY 759 IDTFSDTFDGTGTPISINANINEQSGKDTNNTGNSSETSPVSHPEPESDAAINVEKLSGD 818  
Db 1283 -----KKAEEKKAAEAAKAAEERKRIEAEKAAEERKRIEAEKAAEERKRIEAEK 1321  
QY 819 ESSSETRGILDINDPSVTNNVNEVHDASNTQGSVSNSTSDITNGHSSSSSLNRTTNAQD--- 875  
Db 1322 ESRRRKKRLKLRQKKKKKRLKQRKQKKKKKRLKQRKQKKKKKRLKQRKQKKKKKRLKQR 1381  
QY 876 -IKIGRSGNEQSDNOENSSHSDNSGSLTIQVPSSEDTQNTQYDSQNPRTDTPNALASLP 934  
Db 1382 RIEAEKAAEERKRIEAEKAAEERKRIEAEKAAEERKRIEAEKAAEERKRIEAEKAAE 1440  
QY 935 SDDKINEIGFSSRDSSENGRGTTSNTHDVR-----RTNIVSERVNSHDFI 982  
Db 1441 ABEERKRIEAVKKA--BEERKRIEAEKAAEERKRIEAEKAAEERKRIEAEKAAEERKRI 1499

QY 983 RRGHANNNAHVQITQENNGIIRGQESAGSNVYKNDPKRNFPS-----ENDHK-NI 1037

Db 1500 IE--AAKABEERI-----KAEAVKKSEEVTKKNSNLSETKISNNYBTRNIDDNSFKKLDE 1553

QY 1038 QYNSRDTKVRBEIIKLSKQNCNNNEYSMEYCTYSDERNSSPGCSREERKKLCCQISD 1097

Db 1554 EYKSRNIDTRNKIIISMSKENMCTNDVSSKYCDYMKDKISS-GNCSNDRBKQLCCSISD 1612

QY 1098 YCLKYFNFYSEYNYCJIKSIKPEYKCPKSEGOSSIPYPAAGGILVIVILL 1150

Db 1613 YCLNFYDYNKYYDCTKKFSDPLYKCFNERYSKAVYPAGAGIIMSILIAI 1665

RESULT 13

C71625

variant-specific surface protein 1 homolog PFB0010w - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 09-Jul-2004

C:Accession: C71625

R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.; Perle, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O. Science 282, 1126-1132, 1998

A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A:Reference number: A71600; MUID:59021743; PMID:9804551

A:Accession: C71625

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1711 <R>

A:Cross-references: UNIPROT:O96108; UNIPARC:UPI0000007EE17; GB:AF001366; GB:AF001362; NID

A:Experimental source: clone 3D7

C:Genetics:

A:Gene: PFB0010w

Qy	346	-----KKYE-----	349
Db	514	PPENTTKIPKLTAEKRTGILKKYKFCKNDSGNNGGOIKQWECHEKNDKODGNGDINN	573
Qy	350	-----ACPGGCRLLPC-----	359
Db	574	CIOGDWKTSKNYYPIYSYSPFYGSIIDMLNESIEWRBLKSCINDAKLCKRKGCNCP	633
Qy	360	SSYTSFMKSKTQMEVLNTLYKKCN-----SGVD-----KNNFLNDLFKCNKNKD	404
Db	634	ECYKRWVEKKCBWDKIKEFFRKQKOLLKDIAGMAGELLEFYLENI FLEDMKNANGDPK	693
Qy	405	LDDFPKN--EKEVDDILDCRYATIATIKSPL-----NGPA	436
Db	694	VIEKPKFELIGKENEVEQDPLKTKTTIDDFLEKELNEAKNCVEKNPDNECPKQKAPGDGAA	753
Qy	437	KND-----	439
Db	754	PSDPPREDITHDGEHSDEDEEBEEBEEOQPAEGTEQGESEKSEKVEVQEQTPOKDT	813
Qy	440	-----VDIASOI-----NVNDLRFGCN--YKSNNKSWNC---TGFTNKPFGT	479
Db	814	EKTVPITTPVDCVTKTLADTGSLSNA-ACSLKYVTGKNYGNWCIAISGITSKGDAI	872
Qy	480	CBPPRQTCILGRTYLLHRGHEBDYKHLHGASIYEAQLLKVKYKKEKDENCALCSIIQNSY	539
Db	873	CVPPRTQELCLYYLKELSDDTTQGLREAFIKTAAQETYLLWQYKEDKQNETAS---TE	928
Qy	540	ADLAD--LIKSDIILKD-----YYGKMEENLKNK-----DKK	572
Db	929	LDIDDPQTQUNGGEIPEDFKQMFYTFGVDRDLFLGRYIGNDLCKYNNNITAVFQNGDHI	988
Qy	573	RNEBSLKIFREKWWDKENKNVWMSAVLK--NKECTKD-----YDK	612
Db	989	PNQOKTDRQBEFGYTKDIWKGMCLCAQEGGKTLTETYNYSNVTFNHGLTGTKLNE	1048
Qy	613	FQKIPQFLRFKFWGDDPCBKREKITYSPESFKVBC-----KKQDCDENTCKNKCSE	664
Db	1049	FASRPSFLRMTWEMGQDFC---RERITQILKERCMWYQYNGDKGDDKKEKTEACTY	1105
Qy	665	YKMWIDLKSEYEKQVTKTKQKNKQYNDIEVK-NKEANYVLKESKE-----	713
Db	1106	YKEWLTNWQDNYKKQRYTEVKGTSPYKEDSDVKESYAHGYLRKILKNIICTSGTDIA	1165
Qy	714	-CK-----DVNFDDKJ---FNESPNEVEDMC-----KKCEIKYL-NEI	747
Db	1166	YCNMCGSTTSDSSNNDNIPESLKYPIIEBECTCKDPSPGVEIPEKKVPKVPULPKPP	1225
Qy	748	KYPTKHDIYDIT-----PSDT-----FGDG-----TPISINANINEQQ	782
Db	1226	KLPRQPKERDPTPALKNAMLSSTIMWSIGIGFATFYLYLKKTKSTIDLLRVINIPK	1285
Qy	783	SGKDTNKGNSETSDFVSHESPESDAAINVEK---LSGDE-----SSSETR	825
Db	1286	SDYDI-----PTKLSPNRYIPTYSGYKRGRIYILEGSDGSDSGYTDHYSDDITSSSESE	1339
Qy	826	-GILDIND-----PSVTNNVNEVHDAS--NTQGSVSNT-----SDITNGHSESSL--N	868
Db	1340	YEELDINDIYAPRAPKYKTLIEVVLSPSGNNTTASGNNTPSDTQNDIQNDGTPSSKITON	1399
Qy	869	RTTNAQDIKIGRSGNQENSHSSDNGSLTIGQVPFSEBDNTQ--NTYDSQNPHR---	924
Db	1400	EWNTLKDEFISQVLOSEQPNVPNDYSS-----GDIPL--NTQPNTRYPDNPEKPP	1449
Qy	925	-----DTNALASLPDDKINIEBFGDSRDSSEN-----	953
Db	1450	ITSIHDRDLYSBESYNNVMVNTNNDIPISGKNGTYSIGIDINDSLNSNNVDIYDEVLK	1509
Qy	954	-----GRGDITSNTHDVRRTNIVSERR-----VNSHDFIRNGMANN-----NAHQ	994
Db	1510	RKENELFGTNHTKKNV-----STNSVAKELCGDPIMNQLODLLHKWDRHRDMCKWNKKE	1565
Qy	995	YITOIE-----NNGIIR--GOEBSAGN-----SVNYKONPKRSPFSSEN-----	1031

Db	1566	VLDKLEKWKNDNSGNGINPSGNTPTPSDIPSGKSLDTPSDNNIPSSNKTINTD	1625
Qy	1032	-----DHKNIOEYNSRDTKRVEEIIKLSK	1057
Db	1626	VSIQHMDNPKPINOFTNMDT--ILEDLEKYNE	1656
RESULT 14			
T28626			
variant-specific surface protein 2 - malaria parasite (Plasmodium falciparum)			
C;Species: plasmodium falciparum			
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004			
C;Accession: T28626			
R;Su. X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herrfeldt, J.A.; Peterson, D.S.			
Cell 82, 89-100, 1995			
A;Title: The large diverse gene family var encodes proteins involved in cytoadherence an			
A;Reference number: 220487; PMID:95330813; PMID:7606788			
A;Accession: T28626			
A;Status: preliminary; translated from GB/EMBL/DBJ			
A;Molecule type: DNA			
A;Residues: 1-2664 <SUX>			
A;Cross-references: UNIPROT:Q26033; UNIPARC:UPI000007565C; EMBL:L40609; NID:5886376; PID			
C;Genetics:			
A;Introns: 2197/3			
A;Note: var-2			
Query Match 6.7%; Score 433.5; DB 2; Length 2664;			
Best Local Similarity 18.5%; Pred. No. 7.2e-11;			
Matches 296; Conservative 175; Mismatches 455; Indels 671; Gaps 70;			
Qy	16	LYNVIRINESI-----IGRTLYNQDSSD--ISRVSPELNHNNHNTIY	58
Db	670	LENLINIDELFQDITEAYGNSQIKGIDTLAKKTOAADDATEQKNTIDL-----LF	722
Qy	59	DSDYEDVNN-KLINSFVENKSVKKRSLSPFNNTKSYDVIIPPSYSYR--NDKFNLSN	115
Db	723	EYDSEEAEEKKIQEEOCPKPTKVRNPGYNN---TYDALAGVAOKLOEAKQELDRN	779
Qy	116	EDNS---GNTNSNNFANTSEISGKDNKQVTFIQKRLTHLPACGKKR-----	159
Db	780	DSRSALKANASQGYKSNQGDPPDFKN-----LCGITQKHSNAIGDSKNPCN	826
Qy	160	-----SIKWTCRNSEKITVC-----VPDKRIOLCVANF--LNSRLTWEKFKEIF	203
Db	827	NKGKREFNVGKWK---KNGGEVKMSHTDLVLPRRQHCFCTSNLEHLNTKSTGLTSDRAIH	883
Qy	204	LISVNTAKLLYNKNEGD-----PSIFCNELRNSFSDRSSFIDGDM-DFGDN-	251
Db	884	SLLDGV---LLAAKKEGEDIKTKLTENDNRSSICRTWKYSFADIGDIIRGTDLMDINGDA	940
Qy	252	-----TDRVKGYINTKFSDYVKEKNVEKLNLIKKEWEKKNKANLWNHIVN	297
Db	941	TGVQNQLKDIFSKITEELKQHPDKFNDNDKYTNSKHTKLRSDWEANRDQWKAATCP	1000
Qy	298	HK-GNISKECAIIPAE--POINWIKEWENFNLMEKKRLFLNI-----KDK	341
Db	1001	TKNGNI--QCQATPHDDYIPQRLRWVWEAWFCEQSRLEYELLRDCGSGCTTGKCNNDK	1058
Qy	342	CVENKYEACFGGRLPCSSYTSFMKSKTKQMEVLNLYK-----KNKSGVDKNFP	392
Db	1059	CAK-----CDKQCEYKTKIOPWADQWNEISNKYQILYWQAIAAINGTEKSTT	1108
Qy	393	LND-----	395
Db	1109	TKDDKDNVIDFLQKLHEANYGTRGPPPEAHPDRRPRRAATSKSDVYETTAGYIHQEART	1168
Qy	396	-----LFKQKNQNDLDDFPFNEKEYDDLCDCRYT-----	424
Db	1169	RECLGQNVFCNNNGNNEYAFSLTPHEYPKACKCNENKASSPEELGRSDSFDDHQTTPREE	1228
Qy	425	-----ATIIKSLFLNGPAKNDV-----DIASQI	446

Db	1229	DEVHSSEGEDESEDEKEEVEEVEHGADEKAGAVSQPEASPTTKDVVVKPCDIVKEL	1288
Qy	447	--NVNDLR-----GFGCNYKSNNEKSWNC-----TGT	471
Db	1289	FSNVDTLQKACSTKYGPGKNY-----GWCRIPTKTSNDVTGBDGGSRVVRSTPESGS	1342
Qy	472	FTNKPFGTCETPRRQTLCLGRYYLLHRGHEDY-----KEHLG-	510
Db	1343	NSDKNGATCIPRRRLRYVGLQEWANKENTETSVSQGEATEARGSEAPAPGKSSSG	1402
Qy	511	-----ASIVEAQLLKVKYKEDENALCIIIONSADLAD-----	544
Db	1403	KETPSDKLRTAFIESAAVETFFFLWDYKKAEWLAKAEQLONGGLDLVSSGDGDPDPQNK	1462
Qy	545	IIMG---SDIIR-----DY-----YGKKMEENLKNVNDKKNRNEESLKI	582
Db	1463	LLNGVIPPDLRLMFLMTGLDYLHVGNGTSDSGNTGNSNNNIVLEASGNKEDMQIKI	1522
Qy	583	E-----KWDENKENVWVMSAVL-----KNKETC-----	607
Db	1523	EKIEQILPKNGGTPLVPKSSAQTPDKWNEHAESIWKGMICALTYTEKNPDTSGARGDNK	1582
Qy	608	--KD---YDKF-----OKI-----	616
Db	1583	IEKDEVEYKFPGSTADKGTASTPTGTYKTQDYDEKVKLEDTSGAKTPSASDTPLLSD	1642
Qy	617	----POFLRMFKEWGDDFCERKEKIYSPESFVEK-----KK-----KDCDB-	655
Db	1643	FVLRPPFYRLBEGWGFCKERRKRL--KQIKECWGSDGDKKYSGDGEQCDRRDTSNEV	1699
Qy	656	-----NTCKNKSEYKWKIDILKSEYKQVDKYTKDKKQMYDNDIDVKNKENVYLKE	709
Db	1700	SADLEGRSCGNSCRFYKWKIKRKRKEYDQANAYSQKTKY-----EGSKGAG--LND	1751
Qy	710	KSKECKDVNFDDKIFNESNEYEDMCKK---CDEIKYLNEIKYPKTKHDIYDITFSDT	765
Db	1752	HNKE-----FCVLGTCTDAAAFLNRLKNGPKCKD-----	1781
Qy	766	FGDGTPISANANINEOQSGKDTNTGNSETSDSPVSH-EPESDAAINVEK---LSGDSS	821
Db	1782	-----NENGNDINFCTEETPRPAENCKPCSSFKINCNGNCRSGDGD	1826
Qy	822	SETRGILDINDPSVTNNVNEVDHASNTQGSVNTSDITNGHSESSLNRTTNAODIKGRS	881
Db	1827	KE-----KNGGTTITG-----NFNTMGTC--TEDVWVHVSQKNAEFEG-----	1864
Qy	882	GNEQSDNQENSSHSSDNGSLTIGQVPSEDNTQNTYDSQNPHRDTPNALASLPSDDKINE	941
Db	1865	-----DGLDEACENAGIFT--GIRKDEWKCKVGLHICKEKGNG-----AIND	1907
Qy	942	-----IEGFDSSRDS-----ENGRGDTTSTNTHVRRRTNIVSE-----	973
Db	1908	QQIILVRALLKRWVEYFLEDYKKIKKLKPCIEENGSTCINGCN-KKCNRVGEWIKLKK	1966
Qy	974	---RRVNSHDFTN--GMAN-----NNAHHQYITQIENNGIIRGO-----ESAGNSVN	1017
Db	1967	DEWTKIKNHYLEKNKEGDKNVTSLVTNLETTLVTOIAAANDKREQTSLDKLTSLG--CN	2024
Qy	1018	YKDNPKRSNPFSSND-----HKKNIQEYNSRDTK-----RVREEII	1053
Db	2025	CPEN-SRKJGNDGNDIADCMNLKLETKIHECKTQHENSVENSDQHPNCGNPPPEEDL	2083
Qy	1054	KLSKQNKCNNEYSMBEYCTYSDERNSSPGPCSEERKK	1090
Db	2084	LLEENPEVQPP---GFCPTPQQEPPEPDKCGKLEKK	2117

RESULT 15  
H71621  
serine/threonine-specific protein kinase (EC 2.7.1.-) PFB0150c - malaria parasite (Plasm  
C;Species: Plasmodium falciparum  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 31-Dec-2004  
C;Accession: H71621













303 DNI---IEKEKKYKLEVELEBKNQIDLLNKEKEKEKEKEKEKEKEKEYDTLI 360  
221 ---KDPSTFCNE-----LNSPSPFRSSFIQDDMDFGGNTBVK 256  
361 KELADEKISILEKVHSIKVREMDIJEKREHNFLEHEDQLDKNSFV-----KNNQLK 413  
257 GY----INTKPSDYVYKQVKNVLANIKKEMWEK--NKANLWNHIV-----NHKGNM--- 302  
414 VYKCEIKNLTELEKEKELKDIEVSKKEINKLINQLNKEKQILAPKNKHKEIHLK 473  
303 ---SKECALIPABEQINWIKEWENFMEKKRULFNLIKQKVENKKYVEACFGGRL 357  
474 BELKESVKITKIQELQEMWDIKQKELDQKQRYN--AQIESISIELSKKE----- 523  
358 PCSSYTSFMKSKTQMEVLTNLYK-----KNSGVDKNFNLDLPKNN----- 401  
524 --KEYNQI---KNTYIIBEINNLNKEIBETNKBYTNLQNNYTNLMLNNDIHLNGNIKT 578  
402 ---KNDLD-----DPFKNEK-----EYDDLCDCRYTATIIKSFNGP 435  
579 MNTQISTLKNVDHLLNBEQIDKLANKEGTLSKISLNVQIMDLKE-----EKDFLNQ 631  
436 AKNDVDIASQNV-----NDLRGFGCNYKS-----NNEKSWNCTGTFTNKPFTC 480  
632 I---VDSLSQIDLLTRKMEKENQMLEQENKYQEMELLRGNKS---SENILNDBEVC 685  
481 EPPRQTLCLGRTYLLHRGHEEDYK-----HLGASIEYEAQLKYK----- 523  
686 DLKRLKJSLKSEMKWQKEHDKLAELKDDCVRAIREMNEKEDKINLKEEYEDKINTL 745  
524 KEKDENALCSLIONSADLADIIGKS-----DIIDKYVYKQW-----SENILNVKDKRN 574  
746 KQONEDI--NYLKEQNEDEKINTLKEEYEHKINTWKEHYEHKINTLQNEHKINTLNEQN 804  
575 EESLKIIFREKWDRE---NKENVWKMVA-----VLKKNKETCKDY-----D 611  
805 EHKINTMKEEYEDQWNTLNEQNKSLKEEYENKINQINSNNEIKIKDVVNEYIEVD 864  
612 KF-----OKIPOFLR-----WFKWGGDDFCERKEKISYFSP----- 644  
865 KLKVTLEKKQKQFKEINYAHIAKHEKEQILLTEMEELKQORDNKYSDLYEKIKLSI 924  
645 ---KVBC-----KKKQDSE-----NCKKNKSEYKWI 669  
925 CMINIISCCDDIENEDIRIEEYINNKGLEKEVEKEHGRHSFLLKSKEFFKNSI 984  
670 DLK-----KSEYKQV---DKYTKDNKRMVDNIDEV-----NKEAN----- 704  
985 EDKSHELKKGHEKOLLKSKDEIBEKNKIKELANDIKKLQDEILVYKQNAQQVDHKK 1044  
705 --VYLKESKECKDVNFDKIPNESPNHEYDMCKDEIKYLN--IKYPTKHDIYDI- 759  
1045 SWILLKDKSKE---KIKDKENQINVEKBEKDLKKDDIRILNEELVKY---KTLNLYLK 1059  
760 -DTPSDTFDGTPTISINA-NINEQQSGKDTNTGNSSETSPVSHPESDAANVEKLS- 816  
1100 KDPQLQNDLLSKIDINSILINE-----GMCVDKIEHIIIDYDEEINKSRNL 1147  
817 ---GDESSESSTRGLINDPSVTNNVNEVDASN-----TQGSVNTSDITNGHSE--SSL 867  
1148 FOLKNEICSLTTEVMELN-----NKNELIEENKLNVLNVDGKKLKDVEKQKKEIKL 1202  
868 NRTTNAODIKIGRSGNQSD--NOENSHSSDNGSLTIGQVPSEDNTQNTYDSQNPHRD 925  
1203 NKQLT-----KCNKQIDELNEEVEKLANNE----- 1227  
926 TPNALASLPDDKLINEIEGPDSSSENGRGDTTNTHDVRRTHIVSERRVNSHDFTIRNG 985  
1228 -----TELITYSND-----LNNKFMKENYL----- 1248  
986 MANNNAHQVITQIENNGIIRGOESAGNSVNYKDNPKRNFSSENDHKNIQENSRDT 1045

1249 -----MMKLDEN-----EDNIKKMK-SKIDDWEKEI-KYREDEK 1280  
1046 KVVPERIIKLSKQKCNKNVSEMEYCTYSDEKNSPFGSPCSREERKKLCCQISDYCLKY--- 1102  
1281 KRNLENNINLKKQNE-----DMCIKYNEMNIKYG-----DICVKYBEM 1318  
1103 ---PNFYSIEYNG- IKSEIKSPY-----KCFKSEQSSIPYPAAGILVIVILL 1149  
1319 SLTYKETISLKYEQIKVYDEKCSQYDEIRQYDEKCFQYD-EINKY---GALLNINITN 1374  
1150 LSSASRMKGKNEEVDIGESNIEA-----TFEEN-----NYLNLKSLRIFNOEVOETN 1195  
1375 KAVDSKDRNNNEIISVDNKVEGIANLYLKOIFELNEBIIRLKGINKISLILYSNELNEKN 1434  
1196 ISDYSEYNYNEKNMY 1210  
1435 SYDINMKHIQELLLF 1449

RESULT 23  
T14029  
Variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)  
N:Alternate names: erythrocyte membrane binding protein 1 (EMP1)  
C:Species: Plasmodium falciparum  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: T14029  
R:Chen, Q.; Barragan, A.; Fernandez, V.; Sundstrom, A.; Schlichterle, M.; Sahlen, A.; C.  
J. Exp. Med. 187, 15-23, 1998  
A:Title: Identification of Plasmodium falciparum erythrocyte membrane protein 1 (pfEMP1)  
A:Reference number: Z17860; MUID:98080592; PMID:9419207  
A:Accession: T14029  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-2228 <CHE>  
A:Cross-references: UNIPROT:O60991; UNIPARC:UPI000007861D; EMBL:AF003473; NID:G2961467;  
C:Genetics:  
A:Note: FCR3S1.2-vari

Query Match 5.7%; Score 370.5; DB 2; Length 2228;  
Best Local Similarity 20.3%; Pred. No. 2.9e-08;  
Matches 177; Conservative 131; Mismatches 279; Indels 283; Gaps 39;

420 DCRYTATIIKSFNGPAK-----NDVDIASQINVDLRGFCGNYKSNKSWNCTGTPTN 474  
64 ESKYTELEIANSKRNPKCKDGKNDVD---RFSVKEQAGY-----DNKKMKCSNGM-- 111  
475 KFPQTCPPRRQTLCLGRTYLLHRGHEEDYKEHLG---ASIEYAQLLYKYEKQDN- 529  
112 ---TCAPFRRLHLCKNFPNMNSNDSSKAKHDLAEVCMAAKYEGESIKTHYPKYDSKY 167  
530 -----ALCSIIIONSADLADIIGKSDII-----KDYKGMENLKNVK---DK 571  
168 PGSDFPNCTMLARSPADIGDIIRGRDLYLGNKKKKKQKTEKREKLEKLEPKKHIDN 227  
572 KRNEESLKIF-----REKWDENKENVKVM--SAVLKN---KETCK----- 608  
228 LKDEAKRNGEDPNFYKLREDWWTANRETWGANCTSKELDNSSYFRATCNDTGGGP 287  
609 -----DYDK-----FKIPQFLRMFKWGGDDFCERKEKIKY 639  
288 SQTHNKCRCDDKGANAGKPKAGDGVTVIPTVFDYVYPQLRWFEWAEDFCRKKKKKL- 346  
640 SPSPKVECKKKD-----CDEN-----TCNKKSEYKKKIDL 671  
347 --ENLEKQCRGDKSDSEYRYCSRNGYDCEQTI SRKGKVRMGKGTCDCFACHSYENNIDN 404  
672 KSEYKQVDKYTKD-----KKNQWYDNIDEKKNKEANVYLKES----- 711  
405 QRQKQFDKQ-KYTKHEISDGGGRKRAVGGGTKYEGYKSFYEKLNKDGVTDAFLGLLN 463  
712 --KECDVNFDDKIFNESPNYEDEMCKKCDDEIKYLNBEIKYPTKTHDIYDIDTDFSTFGDG 769  
464 NERACKDITGGKI-----NFKEVNSGGVGGVGG 492

[illegible]



A;Residues: 2260-2401 <KEE>  
A;Cross-references: UNIPARC:UPI000017B648; GB:M34281

Query Match 5.5%; Score 359.5; DB 2; Length 2401;  
Best Local Similarity 20.9%; Pred. No. 9.2e-08;  
Matches 271; Conservative 214; Mismatches 505; Indels 309; Gaps 57;

QY 18 NVIRINESIIIGRTLYNRQDRESSDISRVNSPE-----LNNHK-TNIYD--SDYEDVNNK 68  
DB 519 NKFNDSKSLNETKSTSEESYQNTLUKKVDDYIKVCLNTNELITNCHNQTTLKQNLQ  
QY 69 LINSFVENKSVKKRSLSFINNKTSDYDIIPPSVSRNDKFNLSSENEEDSGNTSNFPA 128  
DB 579 NIKTIKETNSIDKIYTDKFN-----ILTDKTELETFTGLSLNHNESNNKELLTYF 631  
QY 129 NTSEISIGDNKQVTFIQKRTHLFACGIKRSIKWICRENSKXITVCVPDKIQLCVANF 188  
DB 632 YDLKANLGKNNMLYKQFNE-----KEKAVEDIKKKVND-INKIVSNIETIVTSY 683  
QY 189 -LNSRLTMEKFKEIFLISVNTAEKLLYNKQKPSIFCNELRNSPSPFRSSFIGDDMD 247  
DB 684 NINEDTEN-EIGKSLIELNTKVLKVKANVTN-----LNEIKEKLKQY-----DFQD 729  
QY 248 FCGNTDRVKGINTKFSDDYK-----EKVNEKLNKIKKWEKKNKANLNMHMI 295  
DB 730 FGKEK-----NIKYPDENKIKNDITLQKIDKSIETLTEIKK-----NSNH-I 773  
QY 296 VNHKGNISKECAIP-----AEBPQINLWIKWENFNL--MEKKRLFLNIDKRC----- 342  
DB 774 DEIKGQIDK-LKKVPNKTWFNEIDPK--EIEKKIENIVEIKDKKKNIYKEIDKLLNEISK 829  
QY 343 VENKKYEACGGCRLPCSSYTSFPMKSKTQMEVLNLYKKNSGVQDKNFWLDFKK--- 399  
DB 830 IENDK-----TSLEKLNINLSYKSGSLNGLFLQQID 860  
QY 400 NNKNDLDDFPKNEKEY--DDLDCRYTATIISKFLNGPAKNDVDI-----ASQINVDNLRG 453  
DB 861 EEKKAETHIKAMEAYIDDLNLIKKSQIEKEMWINMDIKMDIKHEKMLNISHDYKI 920  
QY 454 FGCNYSNNKSNWNCGTFTNPFPGTCPPRRQTLCLGRTYLLHRGHEEDYKHHLLGASI 513  
DB 921 YHTTSKNHEEK-----ISDIRKNSLKIIQDF-SEESYINDIKKEL----- 959  
QY 514 YEAQLLYKYEKDENALCSIIQNSY--ADLADIIGSDIIDKYKKMGLENLKNVOK 571  
DB 960 -EKVLESQNNNTDINOYLSKIEINIYILNKLNIKKIIDKVKE-YTDEIEKNNKKINAEL 1017  
QY 572 KRNEESLKI PREKWMDEKENVKMSAVLKNKTKCDYKFKIQFLRFPKWEKGDGFC 631  
DB 1018 SNSEKIITQLKEN--SSLKESQSKIKSTI-----DDNYV 1049  
QY 632 EKKEKTIYSPESFKVECKKDCDENTCNKCSYKWKWIDLK-----KSYEKQVDK 682  
DB 1050 SECIGNITLNTKIYVNEKN--NINTYFKNAEYQNVSNFNFNEMADTKSQYILNIKK 1106  
QY 683 YTKDKKKMVDNIDEVK--NKEANVYLKESKECKVDNFDKJFNESPNYEDMCKKDE 740  
DB 1107 NNGTNTNDY--NKKELBEKHKKNVYKDEAGNTQETIKKKNKELFEKYEQEVTVLLNKYYA 1164  
QY 741 IKYLNEIKYPKTKH-----DIYDIDTFSDTF--GDGTFPIISINANINEQQSGKD--TSNT 790  
DB 1165 VELKN--KPKTYKNYSQIIEIKDAWNTTISQADKSEKKNKELKNEQRIEBEVAKNK 1222  
QY 791 GNSETS DSPVSHPESDAAINVEKL--SGDESSSETRGI-----LDINDPSVTNNVN 840  
DB 1223 SNKAILDIQLSVFPFKIKFLKIDLRKSDCLCKETKD IETKISLSIDTQETKLIENKN 1282  
QY 841 EVHDSATQGSVNTSIDTNGHSESSLNRTNAQDIKIGRSGHEQSDQEN--SSHSDNS 899  
DB 1283 -----ILNTLEK----LESKNOQKNIEDQK--KELDEVNSKIKNIESNVNQHK 1326  
QY 900 GSLTIGOV-----PSEDNTONTYDSQPHRDTPNALASLSDDKKINEIEGFSRDS 951

DB 1327 KNYEIGIVEKINEIAEKANKDQIESTOKLIIP-----TIKNLISPPKANDLEGIDTWNKL 1380  
QY 952 ENRGDGTTSNTHDVRRTNIVSERRVNSHDFIRNGMANNAHQYITQIENNGIIRGOE-- 1009  
DB 1381 --GKYNTEMN-----IYEEFIKSYDLITHYLETVSKEPITYEQIKNKRITQAQWELL 1430  
QY 1010 ---ESAGNSVNYKDNPKRSFNFSSENDHKNIQEYNSRDYKRVREELIKLSKONKCNNEYS 1066  
DB 1431 TWIKNVNKAQSYLDDTEANEFDRIVTFKFN-----KLANDVNDKFT 1470  
QY 1067 MEYCTYS---DERNSSPGPCSRERKLCQISDYCLK-YFNFYSIEYY----- 1111  
DB 1471 NEYSKVNKGFDNISINNKKSTDENLLNLNQTKEMYANIVSKYISYKYEAENIFI 1530  
QY 1112 -----NCIKSEIKSPE-YKCFKSEGOSSIPYPA--GGILVVIVILLSSASRMGKSNEE 1162  
DB 1531 NIPKLANSLNIQIKSSGIDLPKNINIALPYLDSQKDTLTFTIPSEKTSYTTVKIDS 1590  
QY 1163 Y----DIGESNTEATPEENNYLNLK--SRIFNQEQVQETN 1195  
DB 1591 YNTLLDILKRSOELQKQEOALNLIFFENRLLHDKVQATN 1629

RESULT 27  
E71622  
probable membrane associated protein PFB0125C - malaria parasite (Plasmodium falciparum)  
C;Species: Plasmodium falciparum  
C;Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 09-Jul-2004  
C;Accession: E71622  
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;  
Perteira, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.  
Science 282, 1126-1132, 1998  
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A;Reference number: A71600; MUID:99021743; PMID:9804551  
A;Accession: E71622  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1308 <GAR>  
A;Cross-references: UNIPROT:O96129; UNIPARC:UPI000017B62A; GB:AE001374; GB:AE001362; NID  
A;Experimental source: clone 3D7  
C;Genetics:  
A;Gene: PFB0125C

Query Match 5.5%; Score 358; DB 2; Length 1308;  
Best Local Similarity 18.4%; Pred. No. 5.3e-08;  
Matches 265; Conservative 238; Mismatches 496; Indels 444; Gaps 67;

QY 9 FLIPLFLVNVIRINESIIIGRTLYNR-----ODESSDISR---VNSPELNN- 51  
DB 2 FIFLFFFTNDVNTRMFF---LYNKLGTSSRRKFDLILKEKNADVEKKOVTLNDEKQNV 58  
QY 52 -----NHKTNIIY----DSYEDVN-----NKLINSFVENKS-VKKKRSLSFI 88  
DB 59 EEEYKKNKGVKNEEDNFFVFDDKEINKLKIKEQCCKNEFEINEKGYILNDEVSTI 118  
QY 89 NNTK-KSYDIIIPPSYRNDKFNLSSENEEDSGNTSN--PANTSEISIGK----- 137  
DB 119 NNITSLNLDILHSSDKNVCTSYNIYPSNGNNNNNNNNNNVHNSNNIFVNDSHMEHFDD 178  
QY 138 -----DNKOYTFIQKTHLPACGIKRSIKWICRENSKXITVC--VPDKIQLCVAN 187  
DB 179 ITDEFFKIQOTNFFSQFNT---SFENKQNVNEELMKHTDINI CDKIIDKKNN---CN 232  
QY 188 FLNSRLTMEKFKEIFLISVNTAEKLLYNKQKPSIFCNELRNSPSPFRSSFIGDDMD 247  
DB 233 TSLDLIHDNDLND-----NLN-----IYEDNNKKO-DVISTDLFMLEKNNVKNKFNKEID 282  
QY 248 FCGNTDRVKGINTKFSDDYKKEKKNVKEKLANIKKEWKEKKNKANLNMHMI VNHKGNISKECA 307  
DB 283 VVYDT-----STTPEINNDNNKKNLYNLNMQMSDKLLN-----NNKDD----- 322  
QY 308 IIPAEPPQINLWIKWENFNLMEKKRLFLNIDKVKENKYEACFGCRLPCSSYTSFMK 367

Db 323 -----TFVI---NNKFLISENNILLDKDISFIDRKIES--NKCEDYCVANNNNNN 368  
Qy 368 KSKTQMEVLTLNLYKKK-----NSGVDKNNFLNDLPK----- 398  
Db 369 BRNLSLDLENAYSKDCSRINEDRIYNNFEDMDKISHDAPDFIIPSSFNKEBDNGNEK 428  
Qy 399 -----KNNKNDLDDP-----FKNEKE-YDDLCDCRYTATIIKSP--LNGPAKNDVD 441  
Db 429 YQNVFDSKNLNLINVEDPPFSNFBSEKQNFQNCDSNENLWLNKKFDEHNVEKNEIY 488  
Qy 442 IASQINVDLRGFGCNYK-----SNK-----EKSNCVCTPTFNKPPGCEPPRR 485  
Db 489 EPKNVYENE-----NYDQKDVDESSKFFENNVPFMDKKNVNDIIDSQVGEVGNVDEK 542  
Qy 486 QTLCLGRYLLHRG-----HEBDYKE-----HLIGASIVYEAQLLYKKYK 524  
Db 543 LDKBEETFYVETGYNIGDELPNRFEEIDENYKEVBEKFDKMGNGENFFEE--VEEKYD 600  
Qy 525 EKDNALCSIIQNSYADLADIIKGSDDIIKYGKQMEEN-----LNKVNKDKNRNEESLKI 580  
Db 601 EK-----VGKNIFEE-----VBEKFDKMGRENIFABIEEYDEKMGENIFEE 643  
Qy 591 FREKWDNKNVKNVKNVAVLNK-----ETCKDYDKFKIPQFLWPKWKGDDDFCEK 633  
Db 644 VEENFGEKVGKNIPEEVEBEKFDKMGEPFFDEVEBEKFD--EKWGEF--FFDEVEBEKFD 699  
Qy 634 RKEKI-----YSFESFKVECK-----KDCDENTCKNCSYKKW 668  
Db 700 MGENIFESIPKODVEIBETYSERKWEHDEKIKDYDEPEKIHDE--KKEVEBFFLI 756  
Qy 669 IDLKKSEYE-KQVDKYTKDKNNKMYDN-----IDB---VKNEANVYLKESKECKQVNF 719  
Db 757 ADKKEENEDSNVELINDKNFPENKETEIDEKVSKNNEEDFVYENNETFCEIDFL 816  
Qy 720 DDKIFNESPNEBDMCKCDRI-----KYLNEIKYPTK 753  
Db 817 KREDNDSENE-----KEIDSEIGEVINIKYVHLNKNNSYDDVHILTHDFKNELLIEKN 871  
Qy 754 HD-----IYDIDTF--SDIFGDTPIISIVANI-----NEQSGKDTNTGNSSET 795  
Db 872 VDNI CSDDNIYDGNICDGDNIYDGDNIYDGDNIYDGDNIYDGDNIYDGDNIYDGDNIH 931  
Qy 796 SDSPVSHPEPDAAI-----NVEKL-----SGDESSSETRGILDINDPST 836  
Db 932 GDNISGNI DDDNIYDGNISGDNISGDNISGDNISGDNISGDNISGDNISGDNISGDN 991  
Qy 837 NNVAEVDHAS-NTQGSVNTSDITNGHSESLNRTTNAQDIKIGSGNEQSDNQENSHS 895  
Db 992 DYVEENS DIRFVYDKGGMVNELIGEYSEKYMNNIEDNELVIW-SASVKNDKERLNDN 1050  
Qy 896 SDNGSLTIGQVP-SEDNTQTYDSQNHRTDTPNALASLPDDKINEIEGDFDSSRDSNG 954  
Db 1051 IDLANNISNDYIKNNNEDIKNVHDS-----FSISNKSSELHDIENGILEKSISSN- 1098  
Qy 955 RGDITS-----NTHDVRTNIVSERRVNSHD-----FTRNGMANNNAHOXITQ 998  
Db 1099 --DIKSIEVCVKEENIEHHKNNKKCKGLNNDNNLNDENMYCDISNDIFKNNEYTKHVD 1156  
Qy 999 I-----ENNGIIRQOESAGNSVNYKDNPKRSPFSSNDHKNTQBYNSRDTKRVREBI 1052  
Db 1157 VYTFDNNNSNLIIGEDHECHVSSNFB---YFNNISKNTESNNIL-YEQNDKKTNNNS 1212  
Qy 1053 IKLSQNKCNNEYSMEYCTYSDERNSSPGPCSRBERKKLCCQISDYCLKYNFY--SEY 1110  
Db 1213 VK-----HPMTY-----IKGFEYASDSINF 1232  
Qy 1111 YNCIKSEIKSEYKCFKSEGQSSIPYFAAGGLVVLVILLSSASRWGKSNEEYDIGESNI 1170  
Db 1233 LKALKGLPPLFLKC-----KDMKPYMRLFNVLKVI-----ESNDYN-GRKRI 1275  
Qy 1171 EAT 1173  
Db 1276 KVT 1278

## RESULT 28

T18402

asparagine/aspartate rich protein - malaria parasite (Plasmodium falciparum) (fragments)  
C:Species: Plasmodium falciparum  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T18402  
R:Barale, J.C.; Candelle, D.; Attal-Bonnefoy, G.; Dehoux, P.; Bonnefoy, S.; Ridley, R.;  
Infect. Immun. 65, 3003-3010, 1997  
A:Title: Plasmodium falciparum AAP1, a giant protein containing repeated motifs rich in  
A:Reference number: T18929; MUID:97378065; PMID:9234746  
A:Accession: T18402  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-3844 <BAR>  
A:Cross-references: UNIPROT:Q94648; UNIPARC:UPI000007ABAB; EMBL:Y08926; NID:e1154302; P1;  
C:Genetics:  
A:Gene: aarp1

Query Match 5.5%; Score 357; DB 2; Length 3844;

Best Local Similarity 19.0%; Pred. No. 2e-07;  
Matches 302; Conservative 226; Mismatches 552; Indels 510; Gaps 70;

Qy 9 FLIPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELNHNKTN---IYDSDYEDV 65  
Db 309 FHVMIIRKVCVLI FHVVIILKILNK-----MYNNKNNKELTYNDEY--- 351  
Qy 66 NNKLINSFVENKSVKKSRSISFINNKTYSYDIIPPSYSYRNDKFNLSSENEDSGNT--- 122  
Db 352 ---LINEFIRDFYSIEKHSBEIINQ-----KILQERYENDGNNNKI 390  
Qy 123 -----NSNNFA---NTSEISIGKDNKQYTFIQKTHLPACGKKRSI 161  
Db 391 TLFHFKKFTSNDIYLESNYFKWIMNNSSKLECNKRYIKRET-----IQEKNI 444  
Qy 162 KWI CRE-----NSKIVTCVPDRKIQLCVAN---FLNSRLRTM 196  
Db 445 KIKCNIIYIEKLNKSKYEFNEFLDIKYDKEKNIIYLHPLYIDLYMGDIFFLN--LPDL 502  
Qy 197 EKPKEIFLISVNT-----EAKLLYNKNEGKDPISFCNELRNSFSDFRSSFGDD---M 246  
Db 503 KLLKNILFINIQLTFIVIEIKYGVWRNGPQMIPOVGEYDNSLF-FQNDLIGIQAIIIM 561  
Qy 247 DFGNTDRVKYINTKSDYKE---KNVEKLANIKKEWEKKNKAN-LWNHMI VNHKNI 302  
Db 562 NVLPYENRKYLDSSHVINLPHQIWNITKVQYDEPLKXLPWSKNBYNDVQNDYI--HKAD- 618  
Qy 303 SKECAIIPAEBPQINLWK-----EWSNEN---FLME-----KKRLFLNIDKCKVENK 348  
Db 619 DEICINEKYDVEDTKYIDTSPNFSYNNINQNHVWVEHSEDKKPYNN-KIKYI--KKN 675  
Qy 349 EACPGGCRLLPCSSYTSFM-----KSKTQMEVLN-----LYKK 382  
Db 676 DEFF--BEHMKVYEMSLIYNRSKKSNDNITLPVNNKMGSKGNLDDDDGHVKLLFGI 732  
Qy 383 KNSGVKKNFLNDFKQKNKNDL-----DDFFKNEKEYDDLCD--CRYTATIIKSFN- 433  
Db 733 DDKNVNNNTNNINSINNINNVNNTNNNTYVYKVEFKSICDLCKYCKQLKQDENE 792  
Qy 434 -GPAKNVD---IASQINVDLRGFGCNYKS---NNEKSNWCTGTFTNKF---PGTCEP 482  
Db 793 SNPLLRTEGRLTSDNINFNKLDNNDNIDQLLDNDKRYNKGVSVRKESYNEEDRICEK 852  
Qy 483 PRRTCLIGRTYLLHRGHEEDYKEHLHGASIVYEAQLLYKK--YKEKDNALCSIIQNSYA 540  
Db 853 YKSKSL-----EMNEYLEEFNIEIRNPFILYQLIFKNFRNLNIRKINDMYV 901  
Qy 541 DL-ADIIKSGD-IKYGYGKQMEENLNKV-----NKDKK 572  
Db 902 EIEHHIKMADKIFSDIYTNKSTVDSSKIGHIFQFFYDIIIRIFNELYLYEYSNIPRK 961  
Qy 573 RNBSLSKIFREKWDNKNVKNVAVLNKVKCTCKDYDKFQ-----KIQPFLRW 622

Db 962 TS-----YKTYERKGYHKKILIQLLAS-----KQFYQLELVNVPKQFRHHPSPYEL 1010  
Qy 623 FKWEG-----DDF-----CEK 633  
Db 1011 VNKYGYKTHPRSNVNLKLNKSNWLYDAFWPLSAFKDQFOTANEKCIKEENSSYIGCSR 1070  
Qy 634 RKEKIYFSPFKVECKKDCDENTCKKC-----SEYKWDILKXSEYKQVD 681  
Db 1071 EEDKEYLNFSPR-----KQVNLFFYFKNSCLVSIILLIIFIINDEFQRIEKKKE-ELEKE 1126  
Qy 682 KYTKDKKKMVDNIDEVKN-----EANYLKEKSECK-----DVN 718  
Db 1127 EMKGMKADVTGEDDSLLNQKGGTTNNNTCDGVVINEEANRSGMSONVVTYESGINMN 1186  
Qy 719 FDDKIFNESNEYEDMCKCDEIKYLNKIKYPTKTHDIDIDTFSDTFGDTPIISINANI 778  
Db 1187 NDDVLGERTNN-----DQINYMDD-----NDDDDNDNDNDNDNDNDNDNDND 1233  
Qy 779 NEQSGKDTNTGNSETSDSPVSEHPESDAINVEKLSGDESSSETGILIDINDPSTVNN 838  
Db 1234 NDDNND 1290  
Qy 839 VNEVHDA-SNTQGSVNTSDITNGHSESSLNRTTNAQ-----874  
Db 1291 DNDNDEDSNLHGSSSNMNLINSRDNNSVTNSNINTARNNSYIIVDQNGNETLFA 1350  
Qy 875 DIKIGRSGNE-----QSDNQENSHSDNSGSLTIGQVPSEDN-----912  
Db 1351 EILINRPNNSINMEIIVNVTDDNVGDENNRNRHNLNADGLSELIRSLISENNIINNVID 1410  
Qy 913 -----TQNTYDSQNP--HRDTPNALSL-----PSDDKNEIRGFPSSRDSSENGR 956  
Db 1411 ESFLQNNVGGDNPINDDIRNRGGSIFDDNI SPDNSSSDELMDSTNSNVRINLYG 1470  
Qy 957 DTTIS-----NTHDVRRT-----NIVSERRVNSHDFIR-NGMANN 990  
Db 1471 RTNGLYDLNLFISDNEIMGNLTHNVNTTILIKIILISVAMGSAHPFIELNSFYFN 1530  
Qy 991 AHQYITQIENNGIIRQGESAGNSVNYKONPKRSNFSEN-DHKKNIOEYNSRDTKVR 1049  
Db 1531 AVNFGTDNDSDS---EKNEEIGNEENDKKYIDKKYIDNNNEIEKKHCHDQNKIY 1587  
Qy 1050 EEIKLSKQKCNV-----EYSMEYCTVSDERNSSPGPC--SR-----1086  
Db 1588 DDVRLSSVSTSYDILMNLHLESCMYTSYDETYERKNGVNYNMRGPNVENLSYDKIL 1647  
Qy 1087 ERKGLCCQISDYCLKY--FNFYSI-EYNYCISEIKSPEYKCFKSEGOSSIPYFAAGIL 1143  
Db 1648 NEKDIVEIENEKEKYIKINDHTVSRNYGV--SLLKDEPKFSTDGNN-----1695  
Qy 1144 VVIVLLSSASRMKGSNEVD-----IGESNIEATPEENNYLNKLSRIFNQ 1189  
Db 1696 -----NIINRCNNNNYIDNMKHDVKGGEGEDEGEDEECQIKES--YKMSSECNK 1746  
Qy 1190 E-----VOETNISDSYNYNEKMY 1210  
Db 1747 ENIFDSISVLRKNNIKRLXNCKNKCNY 1776

## RESULT 29

Tl8501  
hypothetical protein C0760c - malaria parasite (Plasmodium falciparum)  
C;Species: Plasmodium falciparum  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: Tl8501  
R;Lawson, D.; Bowman, S.; Barrell, B.  
submitted to the EMBL Data Library, August 1997  
A;Reference number: Z18935  
A;Accession: Tl8501  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-3394 <LAW>

A;Cross-references: UNIPROT:O77384; UNIPARC:UPI0000080686; EMBL:Z98551; NID:el331903; PI  
C;Genetics:  
A;Map position: 3  
A;Note: C0760c

Query Match 5.5%; Score 356.5; DB 2; Length 3394;  
Best Local Similarity 18.5%; Pred. No. 1.8e-07;  
Matches 261; Conservative 244; Mismatches 509; Indels 397; Gaps 63;

Qy 17 YNVIRINESI-IGRTLYNRQDE-----SSDISRVNSPELNNHKTNIYD--SD 61  
Db 1852 YDIIEIKERLKIETKINKQEKYIIQLOKDNLLINDFNSTTTTNNNNNDNDNDND 1911  
Qy 62 YEDVNNKLNSFV-----ENKSVKKSRLSFINNNKTSYD-----IIPPSYSYRND 107  
Db 1912 NNDNNNDTYQOFTHSLKANLENSRLK-ELSNLEKIQLSDKRNMKITILEDKLPKE 1970  
Qy 108 K-----FNSLSENEEDSGNTSNNNFANTSEISIGKDNKQYTFIQKR 148  
Db 1971 KDKQKLOQIIDDNNKNYMIQYNKLTNLDMLSEENRMLLNKEEY-----EKQ---IEQL 2022  
Qy 149 THLFAAGIKRKSIIKWCIRENSEKITVCVDPDKIQLCVANFLNSRLRLETMEKFKEIFLISVN 208  
Db 2023 NHDHKLFISTKN-----NDIQI---IENKLOEQVDQYITINE-----KQIIVHLN 2067  
Qy 209 TEAKLLYNKNEGKPSIFCNELNSPSPRSS--FIGDDM-----DFGNTDR-VKGYI 259  
Db 2068 LQIKLANQNEHRMRSCDIFNVAHSQDNITKNDHMVVGEDIMGDTNHDVNKNIDOGTNOHI 2127  
Qy 260 NTKFSDYY-----KEKNVEKLANIKKEWEKKNANL-----290  
Db 2128 NOGTNOHINOGTNQHDTCDGPNYNYVQVQVATNREDNKNKERNLSOEIYKINENIDLTS 2187  
Qy 291 ----WHTMIVNKHGNIKSKECAIIPAEPEQINLW---IKENWEN-FLMEKKRFLN----I 338  
Db 2188 ELEKKNMLENKYKELKEKEEYIYKLNLDIMLSNNCKKLKESIMMEKVKIIMNNIOE 2247  
Qy 339 KDKCVEN-----KKYEAFCGCGRL---PCSSYTS-FMKSKSTQME 374  
Db 2248 KDBIIEENKXKYNKLDLIDNNYSVVDKSVSCFEDSNIMSPSCNDILNVFNLSKSNKK 2307  
Qy 375 VLTNLKKNKSGVDKXNFNLDLFKKNKNDLDDFPKNEKEYDDLCDCRYTATIIKSPFNG 434  
Db 2308 VCTNM-DICNENMDSISSNNVNNNNNNNNNN--VNNINNNNNI-----N 2350  
Qy 435 PAKNDVDIASQINVDNLRGFCNKGNNKNEK---SWNCTGTFTNKFPGTCPEPRRQTLCL 490  
Db 2351 NVKNIIVDINNLY-VNNLQ-----LNKDNNDNIIIKNLIKLFK-----L 2388  
Qy 491 GRTYLLHRGHEEDYKEHLGASIEAQLKYKEXDEN--ALCSII-----QNSYADLA 543  
Db 2389 GSCYL-----YIINRLKQIQLKQNLQILSLEESIKSLNEFINNLKNENKNELI 2437  
Qy 544 DIIKGSDIIDKYKQKKEEN---LNKVNKDKKGNESLKIIFREKWDENKENVK---V 596  
Db 2438 KINNFEIILK--LKNLQDNESCIONLNNYLNKNEELNKI-----NVKNIFKYKGYI 2487  
Qy 597 MSAVLKNKETCKDYDKF-----KQIPQFLRWFKEWGDDFCBKRKEKIYSPESFKVECK 649  
Db 2488 IHLIQSNVFCIKIFKHFNENKIIDQSIINKLLYLKKSFD-----FMYDSVIOEIR 2538  
Qy 650 KKDCEBNTCKNCKSEYKWDILKXSEYKQVDKYTKDN-----KKMYDNDIDEVK 699  
Db 2539 E---NKNIIINQ-----DFLTDEYFKHIQTFTKCNVLIQGVLSILKDTNNDFFIQ 2587  
Qy 700 NKEANVYLKESKE---CK-----DVNFDDKIFNESNEYEDMCKCDEIKYLNKIEKY 749  
Db 2588 NKQSNQOQNGNHNHNCMIYPDDEINVTADQQIFDGTENVQOSQONEEDYVN--NEEMY 2645  
Qy 750 PKTKHDIYDIDT-----FSDTFGDTPIISINANEQOSGKDTNTGNSETS--DSPVSHE 803  
Db 2646 T----DKMDLNNMNGDDDDDDDDDDDDNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2701









A:Residues: 1-1166 <GAR>  
A:Cross-references: UNIPROT:O96219, UNIPARC:UPI000007D874, GB:AE001407, GB:AE001362, NID  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB0630C

[illegible]

```
Qy      879 GRSG--NEQSDNQENSHSSDNGSLTIGVPSDDNT-----ONTYDSQNPHRDTG 927  
       |||   |||| :| | :| | :| | :| |  
Db      839 GENGTYDENSTYDENCCTY--DKNRTYDENRTYDNKRTYDENRTYDENTYDYD-- 894  
       |||   |||| :| | :| | :| | :| |  
Qy     928 NALASLSPDKINEIEGFSDSRSENGRGDTTSHTHVDVRTNVISRRVN-----S 978  
       +|||+||||:+++++:|||:|||::|:  
Db     895 KSCVFHKDDIIINE-EECEKTKEADH---RVNEDTDIDKLQAELLEKKEKIREEYIQTPK 950  
       |||   |||| :| | :| | :| | :| |  
Qy     979 HDFIRMGMANNAHHQQYTITQIENNKGIRGOESAGSNVNY----KNPKRSNFSSNDHK 1034  
       |||   |||| :| | :| | :| | :| |  
Db     951 SDISINMKLDQMDKCHEY---ENFNHL-DDESTYDDLSDYHDFTDDELEGNCFCSNNVVK 1005  
       |||   |||| :| | :| | :| | :| |  
Qy    1035 KNTOEQY-NRSRDTKVRBHEI-IKLSKONKMNEYSMETCYYSDDRNS--SPGPCSREERKK 1090  
       |||   |||| :| | :| | :| | :| |  
Db    1006 MNEUNKITYGRNGGLWVENLSLYVLWDIFTNNIS-KYTVHFPEKNEFIVPKAINEEEKR 1064  
       |||   |||| :| | :| | :| | :| |  
Qy    1091 LCCGISDYCLUKYNFYGSIE-----YYNCIKSEI-----KSPRYCKFKSE 1129  
       |||   |||| :| | :| | :| | :| |  
Db    1065 -----RNFEYNIQSWMPIYINCISIINVICRFPLFKHPILPPKVVK----- 1109  
       |||   |||| :| | :| | :| | :| |  
Qy    1130 QGSSIPYPAAAGHGILWVTL-----LIASSARMG-KNSEERYDIGESNIEATFFENNVYL 1180  
       |||   |||| :| | :| | :| | :| |  
Db    1110 -----SIICVTIAAIKLKHKPHILSPSEMNLAKRAEDYLLIENKDQ--EE--L 1153  
       |||   |||| :| | :| | :| | :| |  
  
Qy    1181 NKLSRIFPQG 1189  
       |||   |||| :| | :| | :| | :| |  
Db    1154 NELCVLPFG 1162  
       |||   |||| :| | :| | :| | :| |
```

RESULT 36

```
rhoptry protein - Plasmodium yoelii
C:Species: Plasmodium yoelii
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28677; C45521
R:Keen, J.; Sinha, K.; Brown, K.; Holder, A.  
Mol. Biochem. Parasitol. 65, 171-177, 1994
A>Title: A gene coding for a high molecular mass rhoptry protein of Plasmodium y
A:Reference number: Z20508; PMID:95021522; PMID:7935623
A:Accession: T28677
A>Status: preliminary; translated from GB/EMBL/DDBB
A:Molecule type: DNA
A:Residues: 1-2269 <KE>
A:Cross-references: UNIPROT:Q62223; UNIPARC:UPI0000080871; EMBL:L27838; MID:g45<
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.  
Mol. Biochem. Parasitol. 42, 241-246, 1990
A>Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Mu. I
A:Reference number: A45521; MUID:91101660; PMID:2270106
A:Accession: C45521
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 2131-2269 <KE2>
A:Cross-references: UNIPARC:UPI000017B646; GB:M34283
```

Query Match	5.1%;	Score 328.5;	DB 2;	Length 2269;
Best Local Similarity	21.8%;	Pred. No. 1.8e-06;		
Matches 315;	Conservative 203;	Mismatches 471;	Indels 457;	Gaps 83;
QY	16	LYNVIR-----INESIIGRTLYN-----RQDESSDISRVNSPELNHNHKTNIYDSYDYE- 63		
Db	167	LYNTIKSYDQIYEGDID-TFYNELSSIVKEDPIDDIEDKTKLENLRSKIDNVYDKIQM 225		
QY	64	-----DVNNKLNLSFVY-----NKSVK-----KKESLSPINNKTKS 94		
Db	226	EIETVKSLNNIETNNKLPNTLIEIKYIYDEISKELNMLEDFPKYKELKS-----NKISD 282		
QY	95	YD-----IIPPSYSYR-----NDKFNLSIENEDNSGN-----TNSNFPAN- 129		
Db	283	YDKKREQLSEYKMLIEIRHNYSQTNVDNKTKEEAAQYDKSNEHMTTIPITNEDEISKI 342		
QY	130	TSEISIGKD---NKQYTFIOKRTFLFACGIIKRSIKWICRENS-----EKITVCVCPDRK 180		



Qy 433 NGPARN-----DVIDASQINVDLRGFCGNYKSNNEK--SWNCTGTFTNKFP 477  
Db 1865 NKIKKMQSKILNKYKQVIEFPRLSKKKNKNELLNGKHVEKKNKTALLNG-- 1922  
Qy 478 GTCBPFRQTLCLGRTYLLHRGHEDYK-----EHLGASIEYAQLLKYYKEKDENALCS 533  
Db 1923 -----GHTFVEDQKGEYKKEBERHIVQGEIKEKE--KYTLGGRGSRRS 1967  
Qy 534 IIONSY-----ADLADIKGSDIIKQYVKKMBENLNKVNKKRNEESLK----- 579  
Db 1968 KESUSFRGRGSRRSKEVDTLKGR--RDSLKGRKESL-----KGRERDSLKGRKRD 2019  
Qy 580 IFREKWDENKENYKVMASVAVLNKNETCK--DYDKFQKIPQFLRWFKEWG-----DDF 630  
Db 2020 SFRGKERDSFRG-----KERETLKGREDSLKGREDSLKGREDSFRGKERDSF 2069  
Qy 631 CEKKEKIEYSPESFKVECKKDC-----DENTCKKCEB-YK-KWIDLKKSIEYKQVDKYT 684  
Db 2070 RGERDSFRGKERDSFRGKERDTRFRSDRGSRFRNKGDTGVYKSRDINLYKEENKCKDHY 2129  
Qy 685 KDK-----NKKMYDNID-EVKKEANVYLKESKECKOV 717  
Db 2130 VDKHYINKYPEKYSRKFNFNHSVGSVHNAQKYDSLRYEKEKPYKITENKKNEGNEI 2189  
Qy 718 NFDDKI FNESPNEYEDMCKCDEIKYLNIEIKYPTKGDIDYDIDTFSDTFGDGTISINAN 777  
Db 2190 LKYSIENEKKNYD-----KEQNE-----NCILDKDT-----QCN 2220  
Qy 778 INEQSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSSETRGILINDPVTN 837  
Db 2221 VNT-----KEKNLNDKKSFPSPNIK-----VKLEBEKSDDKR-----DDKND 2259  
Qy 838 NVNVEHDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSGNSQSDNQHSSSD 897  
Db 2260 NTRKNLNDKKSFPSPN-----IKVLEBEKSDDKRDK----- 2294  
Qy 898 NSGLITIGQVPESDNT--QNTYDSQNPDRTPNALASLPDDKI--NEIEGDFSRSRDSN 953  
Db 2295 -----KNDNTRKNLNDK-----SFPSPNIKVLEBEKSDDKRDK 2333  
Qy 954 GRGDTTINTHDVRR--TNI--VSERRVNSHPDRNGMANNNAHQYITQIENN----- 1002  
Db 2334 NDNTRKNLNDKKSFPSPNIKVKREBEKSDKMDKDKKKNEN--TREKNLNDKCLP 2391  
Qy 1003 -GIIRGQESAGNSVYKONPKRNFSPSENDHKNIQEYNSRDTKRVREIILKS--KON 1059  
Db 2392 PSNIKVKLEBEKSDKMDKDKKDKONKKDKKDKKDKKDKKDKKDKKDKKDKK 2451  
Qy 1060 KKN--BYSMEYCTYSDERNSSPGPCSRERKLCQOISDYCLKYFNFYSIERYNCIKSEI 1118  
Db 2452 KKNKKVEKKEIKNDKKN-----NKNKKVA-----KKNKIKKEIKDEI 2492  
Qy 1119 KSPRYK 1125  
Db 2493 KDNKKC 2499

RESULT 38  
A45605  
A;Title: Repeat structures in a Plasmodium falciparum  
C;Species: Plasmodium falciparum  
C;Date: 22-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C;Accession: A45605; A54517  
R;Coppel, R.L.  
Mol. Biochem. Parasitol. 50, 335-347, 1992  
A;Reference number: A45605; MUID:92158014; PMID:1741020  
A;Accession: A45605  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1526 <COP>  
A;Cross-references: UNIPROT:P08116; UNIPARC:UPI0000177F9B

A;Experimental source: Papua New Guinean isolate FC27  
A;Note: sequence extracted from NCBI backbone (NCBIN:83648, NCBI:83656)  
R;Coppel, R.L.; Culvenor, J.G.; Bianco, A.E.; Crewther, P.E.; Stahl, H.D.; Brown, G.V.;  
Mol. Biochem. Parasitol. 20, 265-277, 1986  
A;Title: Variable antigen associated with the surface of erythrocytes infected with malar  
A;Reference number: A54517; MUID:87014571; PMID:3531849  
A;Accession: A54517  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 222-443, 'K', 445 <COP>  
A;Cross-references: UNIPARC:UPI000002C732; GB:M15319; NID:g160060; PID:g552170  
C;Superfamily: ring-infected erythrocyte surface antigen; dna amino-terminal homology  
C;Keywords: surface antigen; tandem repeat

Query Match 5.1%; Score 328; DB 2; Length 1526;  
Best Local Similarity 20.0%; Pred. No. 1.2e-06;  
Matches 270; Conservative 203; Mismatches 474; Indels 402; Gaps 59;

Qy 46 SPELNHNHTNIYDSYEDYVNNKLINSFVENKSVKKRSLSLFINNKTKSYDIIPPSYSYR 105  
Db 316 SKETGESKETRIYE--ETKYNKITSEFRETEVNVK-----ITEESKDRGNKVGSGYE 365  
Qy 106 N-----DKFNSLSNEDNSGNTNSNNTFANTSEISGKDNKQVTFIOKRTHLFACG 155  
Db 366 NSENSNVTSEETKLAKEENEGEKLGEN-VNDGASENSEDPKLQTEEN-----G 418  
Qy 156 IKRSIKW--ICRENSEKITVCPDRK-----IQLCVANFL-NSRLETME----- 197  
Db 419 TKESSEBTDDKPRENEKADNKKSKKKKSPFQMLGCMFLCNKNIETDDEETLVVKD 478  
Qy 198 --KFKIEFLISVNTAEAKLYNQEGKDPISFCNLRNSFSDFRSSFIDGDDMDFGNTDRV 255  
Db 479 DAKKKHFLREANTEK-----NDNEKKDKLL--GE-----GDKEDVKEKDEQ 519  
Qy 256 KG-YINTKFDYDYKQVKEKLNLI-----KKEWEKKKANLWNMI-VNHKNISKECAIL 309  
Db 520 KDKVLGSGDEKEDVKEKNDQKQVLGSGDEKEDVKEKND-----GKKDKVI 564  
Qy 310 PAEPQINLWIKENNFLEKRLFLNIDKQCVENKYEACFGGRLPCSSYTSFMMKS 369  
Db 565 GSEKQ-----KEIKEK--VEKR-----VKKKCKKKV-----KGIKEN 596  
Qy 370 KQ-----MEVLNLYKKNQSGVDKKN-----FLNDLFGKKNQNDLD 406  
Db 597 DTEGNDVKGPETIIIEVEKEIKQVEDGKENDTEGNDVKGPETIIIEVEKEIKQVE 656  
Qy 407 DFPK-NEKEVDD-----LCRYTATIIKSLFNGPKNQVD-----IASQINVN 449  
Db 657 DGKENDTEGNDVKGPETIIIEVEKEIKQVEDGKENDTEGNDVKGPETIIIEVEKEE 716  
Qy 450 DLRGFGCNYKSNNEKSNW-----CTGTFTNKPFGTCEPPROT-----LCLGRTYLL 496  
Db 717 IKQVEGKENDTESKDVIGQETIIIEVEKEIKQVEKKNENILEIKDIVIGQEVII 776  
Qy 497 -----HRGHEEDY--KEHLIGASTIYEQAQLKYYKEKDENALCSIIQNSYADL 542  
Db 777 EEVKVKVKKVEKIKENHTESKDVIGQETIIIEVEV-----KKEIKQVEGKENDTES 831  
Qy 543 ADIIGSDIIKQYVKKQKQENLNKVNKD-----KKNNEESLKIIFREKWDENKEN 592  
Db 832 KDKVIGSEVIGKGVNBSGPNKDKVTQKEKVEKVEKVKKKVKRVRKRNK--NERKDN 889  
Qy 593 V--WKVMSAVLNKNETCKDVKQKIQFLRWFKEWGDDFCEKKEKIEYSPESFKVECKK 650  
Db 890 VIGKEIMKEDVNEKDTA--NKKKEIEQ-----EKEKEVEKVEKVEKEEV 933  
Qy 651 KDCDENTCKKCKEYKQKWIIDLKSEYKQVDKYTKDK--NKKMYDNIDEVKNKEA----- 703  
Db 934 KEKEE--VKEK-EVKEKEVEKVEKVEKEDTESKDEIKQVEKEVEKVEKEDTENKD 990  
Qy 704 ----NVLKESKEC-----KDNVFD-----KIFNESPNEY----- 732  
Db 991 KVIGQETIIIEIKVEKVKRVRKRNKKNENKDNVIGQIMNEDVNEKDTANKDKVIOEKE 1050





Search completed: November 21, 2005, 20:34:43  
Job time : 90 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:34:49 ; Search time 179 seconds  
(without alignments)  
2824.433 Million cell updates/sec

Title: US-10-677-980-2  
Perfect score: 6481  
Sequence: 1 MKGYFNIYFLIPLFLYNYI.....VQETNISDYSEYNNKMY 1210

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 4178326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pcp.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
2	6481	100.0	1210	5	US-10-677-980-2
3	6081	93.8	1143	3	US-09-924-154-14
4	1557	24.0	1435	4	US-10-153-273-4
5	1537.5	23.7	1421	3	US-09-924-154-13
6	1200.5	18.5	616	4	US-10-293-913A-4
7	1199.5	18.5	616	4	US-10-293-913A-2
8	1112	17.2	1086	3	US-09-924-154-15
9	1088	16.8	1568	5	US-10-712-533A-12
10	993.5	15.3	1501	3	US-09-924-154-17
11	761.5	11.7	749	4	US-10-153-273-6
12	754	11.6	1115	4	US-10-153-273-2
13	750.5	11.6	972	3	US-09-924-154-16
14	601.5	9.3	3542	4	US-10-087-013-2
15	495.5	7.6	2710	4	US-10-153-273-12
16	411	6.3	1985	5	US-10-732-923-3351
17	382	5.9	6761	5	US-10-732-923-15035
18	370.5	5.7	1527	5	US-10-732-923-3354
19	363	5.6	407	4	US-10-087-013-8
20	360.5	5.6	921	4	US-10-153-273-8
21	353	5.4	4226	5	US-10-732-923-22586
22	352	5.4	4226	5	US-10-732-923-22707
23	349.5	5.4	2548	5	US-10-732-923-15009
24	348	5.4	294	4	US-10-087-013-7
25	338	5.2	2719	5	US-10-732-923-8668
26	334.5	5.2	2133	5	US-10-732-923-15030
27	333.5	5.1	351	4	US-10-087-013-9

28	331.5	5.1	2910	5	US-10-732-923-3342	Sequence 3342, Ap
29	329	5.1	3124	5	US-10-732-923-22709	Sequence 22709, A
30	327.5	5.1	308	4	US-10-087-013-11	Sequence 11, Appl
31	326.5	5.0	3127	5	US-10-732-923-22588	Sequence 22588, A
32	325	5.0	700	4	US-10-153-273-10	Sequence 10, Appl
33	323.5	5.0	1478	5	US-10-732-923-3353	Sequence 3353, Ap
34	319.5	4.9	311	4	US-10-087-013-10	Sequence 10, Appl
35	318.5	4.9	1939	5	US-10-732-923-3340	Sequence 3340, Ap
36	318	4.9	2110	5	US-10-732-923-8666	Sequence 8666, Ap
37	310.5	4.8	1441	5	US-10-732-923-3352	Sequence 3352, Ap
38	308	4.8	2184	4	US-10-304-095-6	Sequence 6, Appl
39	301.5	4.7	277	4	US-10-153-273-15	Sequence 15, Appl
40	295	4.6	1647	5	US-10-732-923-8314	Sequence 8314, Ap
41	293.5	4.5	4688	4	US-10-282-122A-76865	Sequence 76865, A
42	289.5	4.5	861	3	US-09-820-843A-109	Sequence 109, App
43	287	4.4	1881	4	US-10-032-585-7646	Sequence 7646, App
44	285.5	4.4	282	4	US-10-153-273-16	Sequence 16, Appl
45	285.5	4.4	980	5	US-10-732-923-3343	Sequence 3343, Ap

ALIGNMENTS

RESULT 1  
US-10-677-980-2  
; Sequence 2, Application US/10677980  
; Publication No. US20050239730A1  
; GENERAL INFORMATION:  
; APPLICANT: Mayer, D.C.Ghislaiane  
; APPLICANT: Mailer, Louis H.  
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE  
; FILE REFERENCE: NIH209.001C1  
; CURRENT APPLICATION NUMBER: US/10/677,980  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: PCT/US02/10071  
; PRIOR FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: US 60/281130  
; PRIOR FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1210  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-10-677-980-2

Query Match	100.0%	Score	6481	DB	5	Length	1210
Best Local Similarity	100.0%	Pred. No.	0				
Matches	1210	Conservative	0	Mismatches	0	Indels	0
Qy	1	MKGYNFYFLIPLIPLFLYNYRINESIIGRTLYNRQDESSDISRVNSPELNNHNTIYDS	60				
Db	1	MKGYNFYFLIPLIPLFLYNYRINESIIGRTLYNRQDESSDISRVNSPELNNHNTIYDS	60				
Qy	61	DYEDVNNKLNSFVKNKSVKKSLSFTNNKTSYDIIPPSYSYRNDKFNLSNENSG	120				
Db	61	DYEDVNNKLNSFVKNKSVKKSLSFTNNKTSYDIIPPSYSYRNDKFNLSNENSG	120				
Qy	121	NTNSNFPANTSEISIGKDNQYTFIQKTHLFCAGIKKRSIKWICRENSKITTVCVDPDK	180				
Db	121	NTNSNFPANTSEISIGKDNQYTFIQKTHLFCAGIKKRSIKWICRENSKITTVCVDPDK	180				
Qy	181	IQLCVANFLNSRLTETMEKFEIPLISVNTAKLLYNKNEGKOPISIFCNELRNSPSPRSS	240				
Db	181	IQLCVANFLNSRLTETMEKFEIPLISVNTAKLLYNKNEGKOPISIFCNELRNSPSPRSS	240				
Qy	241	FIGDDMDFGNTDRVKGYINTKESDYKKNVEKLANIKKEWEKKNANLWNHMI VNHKG	300				
Db	241	FIGDDMDFGNTDRVKGYINTKESDYKKNVEKLANIKKEWEKKNANLWNHMI VNHKG	300				
Qy	301	NIKCEALIPAEBPQNLWIKWENFLMEKKLFLNIKDKCVENKKYKACFGGCRLPQS	360				
Db	301	NIKCEALIPAEBPQNLWIKWENFLMEKKLFLNIKDKCVENKKYKACFGGCRLPQS	360				

Db 301 NISKECAIIPAEERQINLWIKENWENFLMEKKRLFLNLIKDCVENKKYKACFGCRLPCS 360  
Qy 361 SYTSFMKSKTQMEVLNLYKKNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDDLCD 420  
Db 361 SYTSFMKSKTQMEVLNLYKKNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDDLCD 420  
Qy 421 CRYTATIIKSLFNGPAKNDVDIASQINVDLRGFGCNYSKNEKSWNCTGTFNKKPFGTC 480  
Db 421 CRYTATIIKSLFNGPAKNDVDIASQINVDLRGFGCNYSKNEKSWNCTGTFNKKPFGTC 480  
Qy 481 EPPRRQTLCLGRVYLLHRGHEEDYKEHLLGASIEYEAQLLYKKEKDNALCSIIQNSYA 540  
Db 481 EPPRRQTLCLGRVYLLHRGHEEDYKEHLLGASIEYEAQLLYKKEKDNALCSIIQNSYA 540  
Qy 541 DLADIIGKSDIIKDYGKMEENLNKVNKKRNEESLKIIFREKWDENKENVKWSAV 600  
Db 541 DLADIIGKSDIIKDYGKMEENLNKVNKKRNEESLKIIFREKWDENKENVKWSAV 600  
Qy 601 LKNETCKDYDKFQKIPOFLRWFKWGGDDFCERKEKIIYSPESPKVECKKCCDENTCKN 660  
Db 601 LKNETCKDYDKFQKIPOFLRWFKWGGDDFCERKEKIIYSPESPKVECKKCCDENTCKN 660  
Qy 661 KCSEYKKWIDLKSEYKQVDKYTKDNKKMYDNIDEVKNKEANVYLKEKSKCKOVNFD 720  
Db 661 KCSEYKKWIDLKSEYKQVDKYTKDNKKMYDNIDEVKNKEANVYLKEKSKCKOVNFD 720  
Qy 721 DKIFNESPNEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDFGDTGPISINANINE 780  
Db 721 DKIFNESPNEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDFGDTGPISINANINE 780  
Qy 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSESTRGILDINDPSVTNNVN 840  
Db 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSESTRGILDINDPSVTNNVN 840  
Qy 841 EVHDASNTQGSVNTSDITNGHSESSLNRTTNAODIKIGRSGNQSDNOENSHSSDMSG 900  
Db 841 EVHDASNTQGSVNTSDITNGHSESSLNRTTNAODIKIGRSGNQSDNOENSHSSDMSG 900  
Qy 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960  
Db 901 SLTIGQVPSDNTQNTYDSQNPHRDTPNALASLPSDDKINEIEGFDSSRDSSENGRDTTS 960  
Qy 961 NTHDVRRTNIVSERRVNSHDFIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020  
Db 961 NTHDVRRTNIVSERRVNSHDFIRGMANNNAHQYITQIENNGIIRGOESAGNSVNYKD 1020  
Qy 1021 NPKSNFSSNDHKNIQEVNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDERNSSP 1080  
Db 1021 NPKSNFSSNDHKNIQEVNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTYSDERNSSP 1080  
Qy 1081 GPCSRERBKLCQIISDYCLKYPNFYSIEYNYNCIKSEIKSPYKCFKSGQSSIPYFAAG 1140  
Db 1081 GPCSRERBKLCQIISDYCLKYPNFYSIEYNYNCIKSEIKSPYKCFKSGQSSIPYFAAG 1140  
Qy 1141 GILVIVLLISSASRMGKSNEYDIGESNIEATFEENNYLNKLSRIFNQEVQETNISDYS 1200  
Db 1141 GILVIVLLISSASRMGKSNEYDIGESNIEATFEENNYLNKLSRIFNQEVQETNISDYS 1200  
Qy 1201 EYNTNEKMY 1210  
Db 1201 EYNTNEKMY 1210

## RESULT 2

US-09-924-154-14  
; Sequence 14, Application US/09924154  
; Patent No. US20020127241A1  
; GENERAL INFORMATION:  
; APPLICANT: Narum, David L.  
; APPLICANT: Sim, Kim L.  
; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use  
; FILE REFERENCE: 05213-0465 43170-262105  
; CURRENT APPLICATION NUMBER: US/09/924,154

; CURRENT FILING DATE: 2001-08-07  
; PRIOR APPLICATION NUMBER: US 60/223,525  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 1143  
; TYPE: PRT  
; ORGANISM: Mammalian  
US-09-924-154-14

Query Match 93.8%; Score 6081; DB 3; Length 1143;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1130; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MKGYFNIYFLIPIFLYINVRINESIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS 60  
Db 7 MKGYFNIYFLIPIFLYINVRINESIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS 66  
Qy 61 DYEDVNNKLINSFVENKSVKKGSLSFNNKTSYDIIIPPSYSYRNDKFNLSSENDNSG 120  
Db 67 DYEDVNNKLINSFVENKSVKKGSLSFNNKTSYDIIIPPSYSYRNDKFNLSSENDNSG 126  
Qy 121 NTNSNFPANTSEISIGKDNKQYTFIOKRTHLFACGIRKRSIKWICRENSKITTVCVDPDRK 180  
Db 127 NTNSNFPANTSEISIGKDNKQYTFIOKRTHLFACGIRKRSIKWICRENSKITTVCVDPDRK 186  
Qy 181 IQLCVAFNLSRLSETEWKEFKEIFLISVNTAEKLLYNKEGKPSIFCNELRNSPSPDRSS 240  
Db 187 IQLCIANFLNLSRLSETEWKEFKEIFLISVNTAEKLLYNKEGKPSIFCNELRNSPSPDRSS 246  
Qy 241 FIGDDMDFGNTRVKGYNITKPSDYKKEKNVEKLANNIKKEWMEKKNKANLWNNHMI VNHKG 300  
Db 247 FIGDDMDFGNTRVKGYNITKPSDYKKEKNVEKLANNIKKEWMEKKNKANLWNNHMI VNHKG 306  
Qy 301 NISKECAIIPAEERQINLWIKENWENFLMEKKRLFLNLIKDCVENKKYKACFGCRLPCS 360  
Db 307 NISKECAIIPAEERQINLWIKENWENFLMEKKRLFLNLIKDCVENKKYKACFGCRLPCS 366  
Qy 361 SYTSFMKSKTQMEVLNLYKKNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDDLCD 420  
Db 367 SYTSFMKSKTQMEVLNLYKKNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDDLCD 426  
Qy 421 CRYTATIIKSLFNGPAKNDVDIASQINVDLRGFGCNYSKNEKSWNCTGTFNKKPFGTC 480  
Db 427 CRYTATIIKSLFNGPAKNDVDIASQINVDLRGFGCNYSKNEKSWNCTGTFNKKPFGTC 486  
Qy 481 EPPRRQTLCLGRVYLLHRGHEEDYKEHLLGASIEYEAQLLYKKEKDNALCSIIQNSYA 540  
Db 487 EPPRRQTLCLGRVYLLHRGHEEDYKEHLLGASIEYEAQLLYKKEKDNALCSIIQNSYA 546  
Qy 541 DLADIIGKSDIIKDYGKMEENLNKVNKKRNEESLKIIFREKWDENKENVKWSAV 600  
Db 547 DLADIIGKSDIIKDYGKMEENLNKVNKKRNEESLKIIFREKWDENKENVKWSAV 606  
Qy 601 LKNETCKDYDKFQKIPOFLRWFKWGGDDFCERKEKIIYSPESPKVECKKCCDENTCKN 660  
Db 607 LKNETCKDYDKFQKIPOFLRWFKWGGDDFCERKEKIIYSPESPKVECKKCCDENTCKN 666  
Qy 661 KCSEYKKWIDLKSEYKQVDKYTKDNKKMYDNIDEVKNKEANVYLKEKSKCKOVNFD 720  
Db 667 KCSEYKKWIDLKSEYKQVDKYTKDNKKMYDNIDEVKNKEANVYLKEKSKCKOVNFD 726  
Qy 721 DKIFNESPNEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDFGDTGPISINANINE 780  
Db 727 DKIFNESPNEDMCKKDEIKYLNEIKYPTKHDIIYDIDTFSDFGDTGPISINANINE 786  
Qy 781 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSESTRGILDINDPSVTNNVN 840  
Db 787 QOSGKDTNTGNSSETSDSPVSHPESDAAINVEKLSGDESSESTRGILDINDPSVTNNVN 846  
Qy 841 EVHDASNTQGSVNTSDITNGHSESSLNRTTNAODIKIGRSGNQSDNOENSHSSDMSG 900



```
Db      985 EEVKEHTSNSDNVQSGGI VNMNVYKELKOTLENPSSSLDEGKAHELSBPNLSSDQMS 1044
Qy      847 NTQGSVNTSDITN---GHSESSLNR-----TTNAQDIKIGRSGNEQSDN-----QENS 892
Db      1045 NTPGLDNTSEHTTERISNNEYKVNEREDRTLTKYEDIVLKGSHMRSDGGLYDENS 1104
Qy      893 SHSDNSGSLTI-CQVPSDNTQNTY-----DSQPHRDTNALA-----SLP- 934
Db      1105 DLSTVNDESEDAEAKMGNDTSEMHSNOSHIESDQDQNDKMTVGLGTTHVQNEISVPV 1164
Qy      935 ---SDDKINE-----IEGFDSSRDSNGRGDTTSTNTHDVR-----R 967
Db      1165 TGEIDELRESKESIKHABERLSHTDIHKINPDRNSWTLHLKDIREENERHLTNQN 1224
Qy      968 TNIVSRVNSHDPIRNGMANNNAHQYITQ-----IENNGIIRGQESAGSNVYKDN- 1021
Db      1225 INISOERDLQKHGF---HTMNNLHGDGVRSQINSHSHGNGRQDRGSGNVLNMSNN 1280
Qy      1022 -----PKRSNFSSNDHKNIQBYNSRDTKRVREBIIKLSKQKCNNEYMEYCTYSDE 1075
Db      1281 NNFNIPSRYNL-----YDKKLDLDIYENRNDSTTKELIKKLAETIKCENBISVKYCDHMH 1337
Qy      1076 RNSSPGCSREERKLCOSDYCLKYFNFYFVSIIEVYNCIKSEIKSPYKCFKSGQSSI 1134
Db      1338 EEPLKCTKEKTRNLCAVSDYCMSTFTYDSEBYNCTKREPDPSYTCRKEAFSSM 1396

RESULT 4
US-09-924-154-13
; Sequence 13, Application US/09924154
; Patent No. US2002012741A1
; GENERAL INFORMATION:
; APPLICANT: Narum, David L.
; TITLE OF INVENTION: Anti-plaasodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT APPLICATION NUMBER: US/09/924,154
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/223,525
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1421
; TYPE: PRT
; ORGANISM: Mammalian
US-09-924-154-13

Query Match      23.7%; Score 1537.5; DB 3; Length 1421;
Best Local Similarity 28.1%; Pred. No. 2.4e-77;
Matches 411; Conservative 211; Mismatches 467; Indels 375; Gaps 44;

Qy      1 MKGVFNIVFLIPLIFLYNIRINESIIGRTLYNRQDESSDISRVNSPELNNHKNITVDS 60
Db      1 MKCNISYIFFPASFVFLVFAKARNEYDI-----KENEKFLDVYKEKFNELDKKKYGVNQKT 55
Qy      61 DYEDVNNKLINSFVENKSVKKRSLSPINN-----KTKSY---DIIPPSYSYRNDKFNSL 112
Db      56 D-----KKIPTPIENK-----LDILNNSKFNKWKSYGTPDNIDKNMSLIN----- 96
Qy      113 SENEDNSGNTNNSNPANTSEISIGKDNKQ-----YTFIQKTH----- 150
Db      97 KHNEEMFNNNYQSFLLSTS--SLIKONKYVPINAVRVSRLSFLDSDRINNNGRNTSSNVEV 154
Qy      151 LFACGIKRSIKWICRENSKIT--VCVPDRKIQLCVANFLNLSLETWEKFEIPLISVNT 209
Db      155 LSNCREKRGKWKDCKKNDKNSRNTVCIPDRRIQICVNLISIIKIYTKETWKDHFIEASKK 214
Qy      210 EAKLLYNKNEKQOPSIFCNELRNSPDSFRSSFGIDMDFGGNTDRVRYGINTKFSDYKX 269
Db      215 ESQLLKKNNDKNYSKFCNDLKNFLDYGHLAWGMDMDFGGYSTKAENKIQEVFKGAHGE 274
Qy      270 KNVEKLNNIKKEWKEKKNKANLWNNHMI VNHKGNISKECAIIPAEEFPQINLMIKENWENFLM 329
```

```
Db      275 ISBHKIKFRKWKWNEFREKLWEAMLSEHKNNIN--NCKNIPQBELQITQIKEMHGSEFL 333
Qy      330 EKRLFLNIDKCKVENKYEACFCGGRCLPCSSYTSFMKSKTQWELVLTNLYKKKNQSVDK 389
Db      334 ERDNRSKLPKSKCKNNTLYEACEKECIDPCMKYRDMIIIRSKFPWHTLSKEYETQK--VPK 391
Qy      390 NNFLNDLFK--KNNKND--LDDFFKN--EKEYDDLCDCRYTATIIKSFNGPAKNDVDIASQ 445
Db      392 ENAENYLIKISEKNDAKVSLLNNDARYSKYCDCKHHTTLVKSVLNGNDNTIKEKREH 451
Qy      446 INVNDLRGCGCNYS--NNEKSNWCTGTPTNKPCTGCEPRRQTLCLGRYLLHRGHEED 503
Db      452 IDLDDFSKGCDCNKNVDVTNTTKWCEKKPKYKSLTKDVCPVPRQBELCIGNTDRIYDKNLLM 511
Qy      504 YKEHLGASIEYAQLKLYKEKENALGSIIONSVADLADIILKSGDIIKDYVKGKMEEN 563
Db      512 IKEHILAIAYESRIILKRYKNKDDKEVCKLINKTFADIRDIIGGTDYWNDSNRKLGVK 571
Qy      564 LNKVNKDKKRNEBSLKI FREKMWDEKNENYKVSALVKNKETCKDYDKFQKIPQFURWF 623
Db      572 INTNSNVLHRNKQNDKLFREDWMKVIKDVWNVISWVFKDKTVCKE--DIIENIPQPFWF 630
Qy      624 KEGWDDFCERKKEKIYSPSPFKYCKKDCDENTCKNKSEYKWDILKSEYKQV--- 680
Db      631 SEWGDDYQDQTKMI---ETLKVCEKKEPCEDDNCNCKNCKNYSKKEWISKKEEYNKQAKY 687
Qy      681 DKYTKDNKKNKYDNIDEVKNKEANVYLKESKECKDVNFDDKI FNESPNEYDMCKKDE 740
Db      688 OEYQKGNYYKY---SEFKSIKEPVLYKYSKCSNLFNFEDEKSELHSHDYKKNCTWCPE 744
Qy      741 IKYL-----NEIKYP----- 750
Db      745 VKDVPISIRNNEQTSQEAVPESBTEIAHRTETRTDERKNOEPANKDLKNPQSVSGENG 804
Qy      751-----KTKHD 755
Db      805 KDLQBEDLGGSRSEDEVTQFVGNHGI PKGEDQTLGKSDAIPNIGEPGETGISTTEESRHE 864
Qy      756-----IYDITFSDTF-----GDGTPPISINAMINEQOSG--KDTSTN----- 790
Db      865 EGHNKQALSTSDPELSDTLQHLHEDTKENDKLPLESSTITSTESGSSDTETPSISEG 924
Qy      791 --GN-----SETSDSPVSHEPESDA--AINVEKLSGD----- 818
Db      925 PKGNEQKRRDDDSLKISVSPENSRPETAOKTSLNLLKLGVDVDSMPKAVIGSSPNDNI 984
Qy      819-----ESSESTRGILDIN----- 831
Db      985 NVTEQDNTSGVNSKPLSDDDVRPDKNHEEYKHEHTSNSDNVQSGGI VNMNVYKELKDTLE 1044
Qy      832 DPSVTNNVNEVH-----DASNTOGSVNTSDITN---GHSESSLNR----- 869
Db      1045 NPSSLDGKAHELSBPNLSSDQMSNTPGPLDNTSEHTTERISNNEYKKNREGERTL 1104
Qy      870 TTNAQDIKIGRSGNEQSDN---QENSSSHSDNSGSLTI--QVPSDNTQNTY----- 917
Db      1105 TKYEDIVLKSWMNRESDDGELYDENSIDLSTVNDESEDAEAKMGNDTSEMHSNOSHIE 1164
Qy      918 -DSQPHRDTNALA-----SLP-----SDDKINE-----IEGFDSSRDS 951
Db      1165 SDQKNDKMTVGLGTTHVQNEISVPVTEIGDEKLRESKESIKHABERLSHTDIHKIN 1224
Qy      952 ENGRGDTTSTNTHDVR-----RTNIVSRRVNSHDPIRNGMANNNAHQYITQ-- 998
Db      1225 PEDRNSNTLHLKDIREENERHLTNQNINISOERDLQKHGF---HTMNNLHGDGVRS 1280
Qy      999 ---IENNGIIRGQESAGSNVYKDN-----PKRSNFSSNDHKNIQBYNSRDTKRV 1048
Db      1281 QINHSHHGNRQDRGSGNVLNMRNNSNNNFNIPSRYNL---YDKKLDLDIYENRNDSTT 1337
Qy      1049 REEIIKLSKONCKNNEYMEYCTYSDBERNSPGCSREERKLCOSDYCLKYFNFYSI 1108
```

Db 1338 KELIKLALINKENISVYCDHMIHBEIPLAKTCTKEKTRNLCAVSDYCMSTYTDSE 1397  
Qy 1109 EYINCISEIKSEIPIKCFKSEQS 1132  
Db 1398 EYINCTKEFDPSYTCFRKEAFS 1421

## RESULT 5

US-10-293-913A-4  
; Sequence 4, Application US/10293913A  
; Publication No. US20040022805A1  
; GENERAL INFORMATION:  
; APPLICANT: Narum, David  
; APPLICANT: Liang, Hong  
; APPLICANT: Fuhrmann, Steve  
; APPLICANT: Sim, B. Kim Lee  
; TITLE OF INVENTION: Synthetic Genes for Malarial Proteins and Methods of Use  
; FILE REFERENCE: 05213-0464 (43170-280206)  
; CURRENT APPLICATION NUMBER: US/10/293,913A  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/345,051  
; PRIOR FILING DATE: 2001-11-09  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic EBA-175 RII  
US-10-293-913A-4

Query Match 18.5%; Score 1200.5; DB 4; Length 616;  
Best Local Similarity 38.4%; Pred. No. 6.3e-59;  
Matches 246; Conservative 111; Mismatches 236; Indels 47; Gaps 13;  
Qy 154 CGIRKRSIKWT CRENSEKIT--VCVPDRKIQLCVANFLNSRLTMEKFEIPLISVNTAEK 212  
Db 14 CREKRKGKMDCKKNDNSVVCIPDRRIQLCIVNLAIKTYTETWKDHFIEASKESQ 73  
Qy 213 LLYNKREGKDSI ECLNELNSPSPERSFIDGDMDFGNTDRVKGYNITKFSDYKKNV 272  
Db 74 LLLKKNNDKNSKFCNDLNSFLDYGLAMGNDMDFGYSYTKAENKIQEVFKGAGHSE 133  
Qy 273 EKLANNIKKWEKKNKANLWNHMI VNHKGNISKECAIIPAEPPQINLWIKENENFLMEKK 332  
Db 134 HKIKNPKKWNNEFKELWEALSEHKNIN--NKNIPQEBLQITQWIKWHGFEFLERD 192  
Qy 333 RLFLNIDKCKVENKKEACFCGCRPLCSSYTSFMKKSQTQMEVLTNLYKKNKSGVDKNNF 392  
Db 193 NRAKLPSKCKNNALYEACEKIDPCMKYRDWIIRSKFEWHTLSKEYETQK--VPKENA 250  
Qy 393 LADLFK--KNKND--LDDFFKN--EKEYDDLCRYTATIIKSPNGPAKNDVDIASQINV 448  
Db 251 ENYLKISEKNDKAVSLLNLCDAEYSKYCDCHHTTLVKSVLNGNDNTIKREHIDL 310  
Qy 449 NDLRGFCGNYSK--NNEKSNWCTGTFTNKGPGTCEPPROTCLGRTVLLHRGHEEDYKE 506  
Db 311 DDFSKFGCDKNSVDNTNKTWCKPKYKSTKDVCPVPRRQELCLGNIDRIYDKNLLMIKE 370  
Qy 507 HLLGASIEAQLLYKYKEDENALCSIIQNSYADLADIIGSDIIKDYKQKMEENLNK 566  
Db 371 HILAIAIYESILKRYKKNDDKEVCKIIQKTFADIRDIIGTDYNDLSNRKLVGKINT 430  
Qy 567 VNKDKRNEESLKI PREKWDENKENVKMSAVLNKNCETKQYDKFKQIPQIRLWPKW 626  
Db 431 NSNTVHRNKQNDKLFREDSEWVKIKDVMNVSIFKDKTVCCK--DDIENIPQFRWFSEW 489  
Qy 627 GDDCEKEKEKIIYSFESKVECKKCDENTCKNKSEYKKNWIDLKSEYKQV---DKY 683  
Db 490 GDDYCDQKTKMI---ETLKVCKEKPCEDDNCKRKNYSKEWISKCKEENYKQAKQY 546  
Qy 684 TKDKNKKMYDNIDVKNKEANVYLKESKCKQVNFDDKIFNESPNEVEDNCKCKDEIKY 743

Db 547 QKGNRYKMY---SEFKSIKPEVLYLKYSKCSNLNFEDEPKELHSDYKNCCTWCPEVK- 602  
Qy 744 LNEIKYPTKTHDIYDIDTFSDFGDTGISINAMINEQOS 783  
Db 603 -----DVPISIRN-NEQTS 616

## RESULT 6

US-10-293-913A-2  
; Sequence 2, Application US/10293913A  
; Publication No. US20040022805A1  
; GENERAL INFORMATION:  
; APPLICANT: Narum, David  
; APPLICANT: Liang, Hong  
; APPLICANT: Fuhrmann, Steve  
; APPLICANT: Sim, B. Kim Lee  
; TITLE OF INVENTION: Synthetic Genes for Malarial Proteins and Methods of Use  
; FILE REFERENCE: 05213-0464 (43170-280206)  
; CURRENT APPLICATION NUMBER: US/10/293,913A  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/345,051  
; PRIOR FILING DATE: 2001-11-09  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic EBA-175 RII  
US-10-293-913A-2

Query Match 18.5%; Score 1199.5; DB 4; Length 616;  
Best Local Similarity 37.4%; Pred. No. 7.2e-59;  
Matches 252; Conservative 113; Mismatches 239; Indels 69; Gaps 15;  
Qy 121 NTNSNNTANTSEISIGKDNQYTFIQKTHLFCAGIKRKSIIKWCIRENSEKIT--VCVPDR 179  
Db 3 NTSSNN-----EV-----LSNCREKRKGKMDCKKNDNSVVCIPDR 40  
Qy 180 KIQLCVANFLNSRLTMEKFEIPLISVNTAEKLLYNKNEGKOPISFNEELNSFSDFRS 239  
Db 41 RIQLCIVNLNSIIKTYTETWKDHFIEASKESQLLKKNNDKNSKFCNDLNSFLDYGH 100  
Qy 240 SPFGDDMDFGNTDRVKGYNITKFSDYKKNVEKLNLIKKEWKNKANLWNHMI VNHK 299  
Db 101 LAMGNDMDFGYSYTKAENKIQEVFKGAGHSEHKIKNFRCKWNNEFKELWEALSEHK 160  
Qy 300 GNISKECAIIPAEPPQINLWIKENENFLMEKKRLFLNIDKCKVENKKEACFCGCRPLC 359  
Db 161 NNIN--NKNIPQEBLQITQWIKWHGFEFLERDNRKLPKSKCNNTLYEACEKIDPC 219  
Qy 360 SSYTSFMKKSQTQMEVLTNLYKKNKSGVDKNNFLNLFK--KNKND--LDDFFKN--EKEY 415  
Db 220 MKYRDWIIRSKFEWHTLSKEYETQK--VPKENAENYLIKISEKNDKAVSLLNLCDAEY 277  
Qy 416 DDLCDCEYATIIKSPNGPAKNDVDIASQINVNDLGRFCGNYSK--NNEKSNWCTGTFT 473  
Db 278 SKYCDCHHTTLVKSVLNGNDNTIKREHIDLDDDFSKFGCDKNSVDNTNKTWCKPKYK 337  
Qy 474 NKPGTCEPPROTCLGRTVLLHRGHEEDYKEHLGASIEAQLLYKYKEDENALCS 533  
Db 338 LSTKDVCPVPRRQELCLGNIDRIYDKNLLMIKEHILAIAIYESILKRYKKNDDKEVCK 397  
Qy 534 IIQNSYADLADIIGSDIIKDYKQKMEENLNKVNKDKRNEESLKI PREKWDENKENV 593  
Db 398 IINKTFADIRDIIGTDYNDLSNRKLVGKINTNSVHRNKQNDKLFREDSEWVKIKD 457  
Qy 594 WKMSAVLNKNCETKQYDKFKQIPQIRLWPKWGDDEKCKEKEKIIYSFESKVECKKDC 653  
Db 458 WNVISWVFKDKTVCCK--DDIENIPQFRWFSEWGGDDYCDQKTKMI---ETLKVCKECKPK 513





```
Db 211 VKNSYLDKDVIFGTDLD-----KNNISKLVESLSKRFK-KOSSVLN--PTAMRRYVGR 263
Qy 290 LWNHMI--VNHKGNISKECAIIPAEBPQJNLWIKW-----NENFLMEKRLFLNAIKDKCVE 344
Db 264 LMKTMIQFYAHLG-----CRKPDENEFQJNRWILEWGKYNCRMLMKEKEKL--LTGCSV 315
Qy 345 NKKEACFGGCLPCSSYTSFPMKKSXTQMEVLNLYKK-----KNSGVDKNLFNDLDF 397
Db 316 NRKSDCSTGCNNECYTYRSLNRQYEVSLGKKYIKVVRYTIFRKRKIQPDNAL--DFL 374
Qy 398 KNNKN---DLDDFFKNE-KEYDDLCDCRYTATIIKSLFNGPAKNDDVDIASOINNVNDR 452
Db 375 KJNCSECKDIDFKPFFFEYKYEKWC-----QSYIDLKIQFQNDI- 418
Qy 453 GFGCNYSKNE-----KSNWCTGFTNKP-----PGTCBPPRQTLCLGR 492
Db 419 ---CSFNAQTDTVSSDKRFLCKEKKPKWCD---KNSFETVHHKGVCSFPRQGFCLGN 472
Qy 493 -TYLLHGRHEEDYKEH-----LLGASIEYAQLLYKYE-KDENALCSLIQNSYADLA 543
Db 473 LNYLL--NDDIYNVHNSQLLIEIMASKQEGKLLWKKHGTILDNQNAKYINDSYDYK 529
Qy 544 DIIKGSIIIDYKGMKEENLN-----KVNKDKRNEESLKI PREKWDENKENVW 594
Db 530 DIVIGNDLWNNISIKVQNNLNLIFERNFGYKVRNKL--FKTIKELKNVWILNRKNVW 587
Qy 595 KVM-----SAVLNKETCKVDYKQKIPQFLWPKEWGDGFCERKKEKIYSFSPKVECK- 649
Db 588 ESMRCGIDEVDQRKTCERIDELENMFQFPWFQMAHFFC---KEKEYWELKNDKCTG 644
Qy 650 ---KDCDENTCKNCKSYKWIIDLKSEYKQVDKTKDKNKNQYNDIBVRNKEANY 706
Db 645 NNGSLCQDKTCQNVCTNMWYTRKLAYEIQSVKY--DKDRKLF---SLAKDKNVTTF 699
Qy 707 LKEKSECKQVNFDDKIFNESPNEYDMCKCD---EIKYLNIEIKYPTKHDIYDI--- 759
Db 700 LKENAKNSNIDF-TKIFDQDLKFLKERCSCMDQVLEVRNKNEMLSIDSSEDATDISEK 758
Qy 760 -----DTFS 763
Db 759 NGEELYVNHNSVSVASGNKEIEKSDEKPEKAQTNGTLTVRTDKDSRNGKDTAT 818
Qy 764 DT----- 765
Db 819 DTKNPENLKVQEHGTNGETIKEBPPKLPESSETLQSQBLEABAQKQKEEBPKKQEE 878
Qy 766 -----FGDCTPISINANINEQQSGK 785
Db 879 EPKKQBEEXKQEQKQBEEXKQBEEXKQBEEXKQBEEXKQBEEXKQBEEXKQBEEXKQ 938
Qy 786 D-----TSNTGNSSTSPVSH-EPESDAAIN--VEKLSGDDESSESTRGI 827
Db 939 EQNVKSSSPVVPQETTSNGSSQ--DTKISSTEPNNSVVDRTATDSMNLDPKRVNENM 996
Qy 828 LDIN-----DPSVTNNVNEVIDA-----SNTQ-----GSVNTS 856
Db 997 SDPNTNTEPDASLKDCKEVDADAKKELQSTVSRIESNEQVQSTPPEDTPTVEKGVGDKA 1056
Qy 857 DI-----TNGHSESSLN-----RTTN----- 872
Db 1057 EMLTSPHATNSSESGLNPPTDDIKTTDGVVKEQBEILGGESATETSKNLEKPKDVEPS 1116
Qy 873 ---AQDIKIGRSGNEQSD-----NQENS-----HSSDNGSLTIGQ 906
Db 1117 HEISEPVLSGTTGKESELLKSKSIETKGETDPRSNDQEDATDDVENSRRDNNLS--- 1173
Qy 907 VPSEDNTQNTYDSQNP-----HRTPNALASLPDDKDKINEIEGFDSS 948
Db 1174 -NSVDNQSNVLNREDPTASETEVVSPEPDDSRRIITTEVPSTTVK-PPDEKRSERVGKEA 1231
Qy 949 RD-----SENGRG-DTTSNTHDVRRTNIVSER-----RVNSHDFIR 983
```

```
Db 1232 KEIKVEPVVPRAIGCEPMENSVSVQSPNPVEDVEKETLISENGLHNDTHRCNISEKDLID 1291
Qy 984 NGMANNNAHHQYITQIENNG--IIRGOEESAGN-----SVNYKDNPKRSNFSSE----- 1030
Db 1292 IHLRLNEAGSTILDDRRNGEMTEGSESDVGELQEHNFSTQQKDEKDFQDIASDREKEEI 1351
Qy 1031 -----NDH-----KKNIOEYNSRDTKVRREIILK 1054
Db 1352 QKLLNTGHBEDVDLWMDRTEDSMGVSNSHLYYNNLSSEKMEQYNNRDASKDREBILN 1411
Qy 1055 LSKQNKCNNEYSMEYCTYSDERNSSPGCSRERKKLCCQISDYCLKYFNFSYIEYNCI 1114
Db 1412 RSNWNTCSNEHSUKYQYMERNKDILLETCEKXELHLCCEISDYCLKFPFKSIEYFDCT 1471
Qy 1115 KSEIKSPYKCFKSEGOSSI 1134
Db 1472 QKEFDPTTYNCFRKQKQFTSM 1491

RESULT 10
US-10-153-273-6
; Sequence 6, Application US/10153273
; Publication No. US20020169305A1
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; Chitnls, Chetan
; Miller, Louis H.
; Peterson, David S.
; Su, Xin-zhaun
; Wellem, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/153,273
; FILING DATE: 21-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 749 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-153-273-6

Query Match 11.7%; Score 761.5; DB 4; Length 749;
```

[illegible]

RESULT 11

US-10-153-273-2 ; Sequence 2, Application US/10153273 ; Publication No. US20020169305A1

FUNCTION NO. US2002  
GENERAL INFORMATION:

APPLICANT: Sim, Kim L.  
Chitnis, Chetan  
Miller, Louis H.  
Peterson, David S.  
Su, Xin-zhaun  
Wellens, Thomas B.

**TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX**

7 TITLE OF INVENTION: BINDING OF PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach  
STATE: California  
COUNTRY: US  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/153,273  
Filing DATE: 21-May-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/210,288  
Filing DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Fuller, Michael  
REGISTRATION NUMBER: 36,516  
REFERENCE/DOCKET NUMBER: NIH121.1FWDV1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 115 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmodium vivax  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-153-273-2

Query Match 11.6%; Score 754; DB 4; Length 1115;

Query Match	Score	DB #	Length
Best Local Similarity	24.5%	Pred. No. 1.3e-33;	
Matches 296; Conservative 177;		Mismatches 416;	Indels 318; Gaps 57;
Query Match	11.0%	Score 75%;	DB #
Best Local Similarity	24.5%	Pred. No. 1.3e-33;	
Matches 296; Conservative 177;		Mismatches 416;	Indels 318; Gaps 57;

[illegible]

Db 428 VNIIEPOLYRIRWREGRDYVSELPEV---QXLEKCDGKLYNTDKVKCKVPPCQACKSY 484  
Qy 666 KKWIDLKSEYKQVDKYTKDNKKKMDNIDVKNKEANV-----YLKESKECKOVNFD 720  
Db 485 DOWITRKNQDVLNKNFISVNAE-----KVQTAGIVTPVDILKQELDEFNEVAPE 536  
Qy 721 DKLFENSPNEVEDMKCKDEIKYLNELKYPKTKHDIYDIDTFSTFG-DGTPI----- 772  
Db 537 NEI-NKRGDYIELC-----VCSVEEAK-KNTQEVVTVNVDNAKSOATNSNPISQPVDS 589  
Qy 773 -----SINANINEQSGKDTNTGNSETSD-----SPVSH-----PESDAAINV 812  
Db 590 KAEPKPGDSHTGVN-----SGQDSSTTKAVTGCQNGNQTPAESDVORSIALSVSAKV 646  
Qy 813 E-----KLSGD-----ESSSETRGILIDNPSTVNNVNEVHDASNTQGSV----- 852  
Db 647 DPQKSVSKRSDDTASVTGIAEAGKENLGASNSRSESTVEANSFGDDTVNSASIPVVSGE 706  
Qy 853 -----SNTSDITNGHSESSLNRTTNAQDIKIRSG-----NEQSDNQENSHSS 896  
Db 707 NPLVTPYNGLRHSKDNSSDGPASMANPDSNSK-----GETGKGQDNDMAKATKDSNS 762  
Qy 897 DNGSLT-----IGQVPSEDNTQVYDSQNPHRDTPNALASLPDDDKINEIEG 944  
Db 763 DGTSSATGDTTDAVDREINKGVPEDRDKTVGSKDGGEDNSANKDAATVVGEDRIENSA 822  
Qy 945 FDSRD-----SENG-----RGDTTSN---THDVRTNIV 971  
Db 823 GGSNDNRSKNDTEKNGASTPDSKQSDATLAKTESLESTESGDRTTNDTNSLENGKG 882  
Qy 972 SERRVNSHDFRNGMANNHQQVIT-----QIENGLIRGOESAGNSVNYKDNPK 1023  
Db 883 KENDLQKHDFKSNDDTNPNEPNSDQTTDAEGHDSIKND-----KAERREKHNKDTFTKNTN 939  
Qy 1024 RSNPSEN---DHKNIOEYNSRDTKVRBEIILKSKONKNNEYSMEYCTYSDERNSSP 1080  
Db 940 SHLNSNNLSNGKLDIKEKYRDVKATREDIILMSSVRKNNISLEYCNSVEDKISS- 998  
Qy 1081 GPCSRERKKLCCQISDYCLKYFNFYIENYCNISBKSEIKPEYKCFKSEG-QSSIPYFAA 1139  
Db 999 NTSREKSKNLCCSISDFCLNPDVYSVEYLSCKMKEFEDPSYKCFKGGFKIDKTYFAA 1058  
Qy 1140 GGLVWVILLSSASRKGNEEDIGESNIEATFEE-NNYLNKLSRI-----FNQVQE 1193  
Db 1059 AGAL--LILLIASRKMKNDS-----EATNFEEFYCDNIHRPLMPNPNIEHQ 1108  
Qy 1194 TNISDYS 1200  
Db 1109 STPLDYS 1115

RESULT 12  
US-09-924-154-16  
; Sequence 16. Application US/09924154  
; Patent No. US20020127241A1  
; GENERAL INFORMATION:  
; APPLICANT: Narum, David L.  
; APPLICANT: Sim, Kim L.  
; TITLE OF INVENTION: Anti-plasmodium Compositions and Methods of Use  
; FILE REFERENCE: 05213-0465 43170-262105  
; CURRENT APPLICATION NUMBER: US/09/924,154  
; CURRENT FILING DATE: 2001-08-07  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 972  
; TYPE: PRT  
; ORGANISM: Mammalian  
US-09-924-154-16

Query Match 11.6%; Score 750.5; DB 3; Length 972;  
Best Local Similarity 25.8%; Pred. No. 1.7e-33;  
Matches 251; Conservative 143; Mismatches 329; Indels 257; Gaps 44;  
Qy 355 CRLPCCSYTSMFKKSKTQMEVLNLTNLYKK--KNSGVDKQNFNL-----DLFKONKNDDDD 407  
Db 35 CKSECKYKAWIDKKNDFILSEIYLKYNKKSLSYKTAPEYLKQKWDYKELNFFSIFD 94  
Qy 408 FFKNEKEYDDLCRCRYATIIKSFNGPAKNDV-----DIASQINVNDLRG-FGCNYKS 460  
Db 95 QL-NAKYTNK-CICQ-----NNKIENNALYKIEDICNNTKVKISIVGEDYCKEK- 141  
Qy 461 NNEKSNWCTGTFNKPCEPTCEPPRQTLCLGRYTLHRGHEEDYK-----EHLG 510  
Db 142 GNDKIQCINEHTKDPFVCGPPRQQLCLGNL-----DKDEFKNVNDLKKFLNELILG 195  
Qy 511 ASIYEQAQLKYKYEK-----DENALCSIIQNSVADLADIIKGSIIIDYKYGKWEEN 563  
Db 196 IR-DEGKFLIEKRYKNWHENMYLDERA-CYLYNSFDDYKXIIILGKDMWRDPNSIKTENI 253  
Qy 564 L-----NKYNKOKKRNEESLKI PREKMWDEKENENVMKMSA-VLNKNET--C--K 608  
Db 254 LKGNPEGIKANIVSMYPSYADLSLDEFKHWBQNKQLWEAISCEPYKGNHTGVCWMED 313  
Qy 609 DYDKFOKIPOFLAWPKWGDGDFCEKKEKIKYSESPKVEC-----KKKDCDEN--T 657  
Db 314 DNDN-----QYLHWPFREWKNDFC---IDKLKWNVDVIKEPCIDKKVSPKSENPSPDVATV 365  
Qy 658 CKNKSEYKWIIDLKSEYEKQVDKYTKDNKKKMDNIDVKNKEANVILKESKECKOV 717  
Db 366 CNKSCDYDKWIIINKKKEYKQSSKYKRDSS--LFNNV--IQNLKPWEYLSMKCTEC-TC 420  
Qy 718 NFDKXIFNESPNEYDMCKKC-----DEIK-----YLNKIKPKTKHDI----- 756  
Db 421 NLDQTQVYKGYEDICKSTVKPYDPEDIKDEFNEPSLVNPLSLTSQDVTERTVSVD 480  
Qy 757 -----YDIDTFSDFGDTPI 772  
Db 481 DVLSIKENVLDKFPKPGQTQSHVQVGNPRESESKPSGANGREDPSTESSTYNDGV-I 539  
Qy 773 SINANTINEQSGKDTNT-----CNSETSDSPVS-HEPESDAA- 809  
Db 540 TSSSSILG-SSSGRDVSSPVGVGDEHEAKELLPPQKIIDVGTQSDSTLSQHKESSEQ 598  
Qy 810 --INVEKLGSDESSSTRGILDINDPSTVNNVNEVHDASNTQGSVNTSDITNGHSE-- 864  
Db 599 HNLGSSLSRHSNQDEERSI-----ITSDVE--HGTNSLFGSQIQDQETILGESEPLT 649  
Qy 865 -SSLNRTTNAQDIKIRSGNEQSDN---QENSSSHSDNSGSLTIGQVPSD-----NT 913  
Db 650 TSPPEHETSKMDTHAGGKNMEQVRNASVDSSSEMSNGRGLKTKEMKGEVITGITSKND 709  
Qy 914 QNTYDSQNPHRDT--PNALASLPDDDKINEIEGF-----DSSRDSDENGSGDT 958  
Db 710 INLEDSVTSRQNKIENSGDNTQCKEHIINVQMDXHELENPTPTSERGSDVLESEFSKLR 769  
Qy 959 TSNTHDVRETNIVSERRVNSHDFIRNGMANNHQQVITO-----IENNGI 1004  
Db 770 TSHTHDNNRIETTAENNI-----GGLSNVNVHDGRDSQRNRMHNSRSHGSLESDIV 822  
Qy 1005 IRG-----QESAGNSVNYKONPKRSNFSSENHKKKIOEYNSRQTKRVRESIILK 1054  
Db 823 VRGDDISNTEGGBEEBEDANTLKY-----PRNLNNKNSRTYNTIEEYIRYDVNKVADIMR 878  
Qy 1055 LSKQNKCNNEYSMEYCTYSDERNSSPQCSREERKKLCCQISDYCLKYFNFYSTEYNCI 1114  
Db 879 SYKSNECTNNLSNYCS-KLKEKESLNTCTNEDSKRLCCSISDYCMKFFNNSGSHSCM 937  
Qy 1115 KSEIKSPYKCFKSEGOSSI 1134  
Db 938 RKEFSNHAYKCFAGKGFSSM 957

Qy	452	RGFC-NYKSNNEKSWN-CTG-----TFTNKPGTCE-----PP 483
Db	1610	KHGGCTKFTIENDKYTNISSKDKCKGLVKEANTGAIKWQNGKPNYNNLKTEDVLFFS 1669
Qy	484	RRQTLCLGRTYLLHRGH-----BEDYKEHLIGASIEAQLLKYYKKEKDENAL-----531
Db	1670	RLRLICF-----HALDGNNTDPEVKDENGRLKRLMEVAATGYNIGQYYKKEKEKIKTSD 1726
Qy	532	-----CSIIQNSYADLADIIGKSDIIXDYGYGKMEENLNKV-NKD-----KKNNE 576
Db	1727	AHKYSVEVPSCSAMKTSFYDLRDIILIGDNLED-EKQKTEENLKKLFNNGTSGVGKSGDS 1785
Qy	577	SL-----KIFREKWDENKENVWKMVAJLK-----NKEBTCKDYDKFKQIP-----617
Db	1786	TTGNPGSTARKFPWNEKENCKWVAMITCGYKGRDDGNGSNARSDEDLKKGSGVPSSDDY 1845
Qy	618	-----OFLRFKXEWGDDCEKEKXIYSPESPKVECKKCDEN-----TKCN 660
Db	1846	PMGNRDEBTAYOFLRWFAWGSDFC-KHKEK--ELEKLVGACNDVTCGDNEBDRKCKCTD 1902
Qy	661	KCSYKKWIDLKSEYEBQVYTKDKNKNKYDNDIVENKNE-ANVYLKEKSKCKDVNPF 719
Db	1903	ACTQYKKFSEWQYKQIKKYGEND-KISEHPVAKADAREYLDQK-----1954
Qy	720	DDKIFNESPNYEDMCKCDEIKYLNBEIKYPTKHDIYDITFSDTFGDTPISTANIN 779
Db	1955	--KICENKSGDCEYKMWK-----DVSTQRLTDGNSQNPASLD 1990
Qy	780	EQSGKOTNTNGSETSDSP--VSHPEPSDAAINVEKLSGDESSSETRGLINDP-SVT 836
Db	1991	DEP--KEVGEKNCQVPRGPPRVRRRTPTSPRVSLISKATA--SKKEAKTAPPTKQPKQVE 2046
Qy	837	NNVNEVHDASNTOGSVSNTSDITNGHSESSLNRTTNAODIKIGRSQNEQSDNOENSSHSS 896
Db	2047	NLTTEBRAQTRERAAQTRKTS-----TATTTESB--VGTWVKAILSNKPSRGCI 2097
Qy	897	DNSSGSLTIGQVP-----SDEN-----TONTYDS-----919
Db	2098	EGCNPKTYGQYPKWGCIVGKSKENENGICWPPRRKKLCINNTQYLYNETENKRDNDI KEA 2157
Qy	920	-----QNPHRDTPNALASLPD-----DKINEISGFDDSSRDS 951
Db	2158	FIKAAIETQFLWKYIIENPAEBSLONGTIPDEFKRIWYTYGDKDMFFGTDISNDK 2217
Qy	952	ENGRGDTSTNTHDVRRTNIVSSRRVNSHD-----FIRNGMANNNAHQ 994
Db	2218	K-----IIVTVSV--TTILNENKKQDKKDEBLKTFWEKNKFIWEGYIGLTYH- 2269
Qy	995	YITQIENNGIIRQGEBSAGNSVNYKD-----NPKRSNFSSENDHKNIQFYSRDTKRVRE 1050
Db	2270	-LTD-EN-----EKEKIRDNTQYNDMTKLTPSLEBEFVKRPQFLRWFTWEAEFCVKRKE 2321
Qy	1051	EIIKLS---KQNKONNEYSMEVCTYSDERNSSPGPCSRERKKLCCQISDYCLKYFNFYVS 1107
Db	2322	QLLKLEAGCKEYECNG-----SNDGKTQBC---AEACVTYQNF-- 2356
Qy	1108	IEYNYCKISEIKSPKCFKSGBOSSIPYFAAGILWIVLVLSSASRMKSGNEEYDICE 1167
Db	2357	-----JKWKTYERQREKF-----KKDKGKKYKDYKPSTE 2387
Qy	1168	SNIEATFEENNYLN-KLSRI 1186
Db	2388	RDIEKATCAHEYNLMKMLKEL 2407

RESULT 14

US-10-153-273-12

; Sequence 12, Application US/10153273

; Publication No US20020169305A1

; GENERAL INFORMATION:

; APPLICANT: Sim, Kim L.

; Citnise, Chetan

; Miller, Louis H.

RESULT 14  
US-10-153-273-12  
; Sequence 12, Application US/10153273  
; Publication No. US20020169305A1  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; Chitnis, Chetan  
; Miller, Louis H.  
;

Peterson, David S.  
 Su, Xin-zhaun  
 Wellens, Thomas E.  
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
 AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
 NUMBER OF SEQUENCES: 37  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Knobbe Martens Olson & Bear  
 STREET: 620 Newport Center Drive 16th Floor  
 CITY: Newport Beach  
 STATE: California  
 COUNTRY: US  
 ZIP: 92660  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/153,273  
 FILING DATE: 21-May-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/210,288  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fuller, Michael  
 REGISTRATION NUMBER: 36,516  
 REFERENCE/DOCKET NUMBER: NIH121.1FWDV1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 235-8550  
 TELEFAX: (619) 235-0176  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2710 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Plasmodium falciparum  
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
 US-10-153-273-12

Query Match 7.6%; Score 495.5; DB 4; Length 2710;  
 Best Local Similarity 19.5%; Pred. No. 1.1e-18;  
 Matches 268; Conservative 170; Mismatches 482; Indels 453; Gaps 57;  
 112 LSENEGSGTNSNPNANTSEISIGKDNKQYTFIQKTHLFCAGIK-RKSIKWICRENSE 170  
 912 INENYSDNRSGSGPCT-----GKGDH-----GGVRVRIETWSNIEGKK 953  
 171 KIT---VCVPRKIQLCVANFLNRLTETM---EKFKIFILISVNTAKL----- 213  
 954 QTSYKAVFLPFRHEMCTSNLENLDVSGVTKNDKASHLLGDVQLAAKTDAAEIKRYKD 1013  
 214 -----LYNKGEGKDPISFCNELRNSFDSFSSFTGDDM--DFPGNTDRVKYI----- 259  
 1014 QNNIQLTDPIQQKQKQAMCAVRYSFADLGLIIRGRDWDDEKSDTDMETRLITVFKNIK 1073  
 260 -----NTKFSYKKEKQVKEKLNKIKKEWEKQKANKLNHNMVNHKGISKECALIPA 311  
 1074 EKHDGIDKNPKYTGDSEKSPAYK--KLRAADWWEANRHQVWRAMKCATKGI---CPGMPV 1128  
 312 EE--POINLWTKWENFLMEKRLFLNIDKQVE--NKKYEAC-----FGGCRLLPCSS 361  
 1129 DDYIPQRLRWTEWAEWYCKAQSGEYDKLKIICADCMKSGDKCTQGDVDCGKCAACDK 1188  
 362 YTSFMKSKTKQMEVLTNLY-----KKKNSGVDKN-----NFLNDLDFK--- 398  
 1189 YKEBIEKWNEQWRKISDKYNLLYLAQAKTTSTNPGRTVLGDDDPDQYQWVDFLTFPHKASI 1248

399 ----- 398  
 1249 AARVLVKRAAGSPTETAAAAPIPTPYSTAAGYIHOEIGYGCQEQTOFCEKHGATSTSTT 1308  
 399 KNNKNDLDDFFKNKEKYDDLDCRYTATIKSLFNGPAKNDVDIASQINVD--LRGFG- 455  
 1309 KENKE--YTFKQPPPYATACDC-----INRSQTEBPKEENVESACKIVKILEGKNG 1361  
 456 -----CNYKSNNEKSNWCTGTFTNKPFGTCBPRTQTLCLGRTYLLHRGHEBDYK--EH 507  
 1362 RTTVGECNPK-ESYPDWDCQNNIDISHDGAQMPRRQKCL--YYIAHESQTEINIKTDDN 1418  
 508 LIGASIVEA---QLLKVKYKEKDNALCSIIQN-----SYADLADIKGSDI 551  
 1419 LKDAFIKTAATAETFLSQYKSKNDSEAKILDRGLIPSQLRSMWYTFGDRYDCLNTDI 1478  
 552 IKDYYG-KKMEENLNK-VNKDKKRNEESLKIFREKWDENKENVKNVMSAVL-----K 602  
 1479 SKQNDVAKAKDKIGKFFSKGSKSPGSL--RQEWKTNQNGPEIKMGLCALTYVTDTD 1536  
 603 NKETCKD---YDK-----FQKIPQFLRMFKWGGDDFCCKRKEKIYSFESFKV 646  
 1537 NKRKIKNDYSYDKVQSNQGNPSLEBFAAKPQLRWMIENGEEFCAERQKK--ENIKD 1593  
 647 ECKK---KDCDE--NTCKNKCSYKKWIDLKSEYEKQVDKYT-----KDKNKKMYD 693  
 1594 ACNEINSTQOCNDKAGHCNQACRAYQYEVENKCKEFGSGTNNFVLKANVQPDPEYKGYE 1653  
 694 NIDEVKNKEANVYLKEKSKCKDVNPDCKIFNESPNE-----YEDCKCKDEIKYL 744  
 1654 YKGVQPIQNGEYLLQKCDNNKSCMDGNVLSVPKPEKPGKYAHKYPEK-DCYQQKHV 1712  
 745 NEIKYPK-----TKHDIDYDITFSDT---FGDGTPIISANINAEQ 781  
 1713 PSIPPPPPVPOPEAPTVDVCSIVKTLFKDTNNFSDACGLKYGTAFSSWKCPISDT 1772  
 782 QSGKDTSTNGSETSDSPV-----SHEPESD----- 807  
 1773 KSGAG-ATTGKSGSDSGSICIPRRRLRYVVKLQEWATALPOGEGAAPSHRADDLRNF 1831  
 808 ---AAINV-----EKLSDGESSESTRGIL---DINPSTVNNVNEVHDSANTQ 849  
 1832 IQSAAIETFLWDRYKEEKKPGDGSQOALSQLTSTYSDEEDPP-----DKLLQN 1882  
 850 GSV-----SNTSDI--TNHSESSLARTTNAQDIKIGRSGNEQ 885  
 1883 GKIPDPFLRLMFYTLGDRDILVHGNTSDSGNTNG-----SNNNNIVLEASNGKE 1933  
 886 SDN--QENSSHSDNSGSLTI-----GOVPSE---DNTQN-----TYDSONPHRDT 926  
 1934 DMQIKQEKIQILPKNGGTPLVPKSSAQTPDKWNEHAESIWKGMI CALITYTEKNPDTSA 1993  
 927 PNALASLPDDKINEIEGFDSSRSDSENGRDTTSTNHDVRRITNIVSERRVNSHDFIRNGM 986  
 1994 RGDENKIEKDEEV--YEKFGSTADKHGTASTTGT----- 2027  
 987 ANNAHQQYTQIENNGIIRGOEESAGNSVNYKDNPKRNSFNSSNDHKKNIQEYNSRDTK 1046  
 2028 -----YKQYDYEVKLEDTSGAKTPSASDTPLLSDFLRPPYFRVLEENGQNFCK 2079  
 1047 RVREBIIKLSKQNK-----CNNEYSMEYCTYSD-ERNSSSPGP 1082  
 2080 KRKHQAQIKHEKVEENGSGSRGGITROYSGDEACNEMLPKNDGTFVDPLEKPSCAKP 2139  
 1083 CSR-----BERKKLCCQISDYCLKYFNFYFISYNYNCIKSIKISPEY 1123  
 2140 CSSYRWKIESKGEFEKQKAYEQKDKCKVNGSKHNDNGFCEILTTSKAKDF 2192

RESULT 15  
 US-10-732-923-3351  
 ; Sequence 3351, Application US/10732923  
 ; Publication No. US20050108791A1











Qy	436	AKQNDVDTASQINVADLRGFGQNYKSNNEKSNWCTGTITNKPGETCBEPRTQTLCLGRTYL	495
Db	1265	MKNRNDIIFETST-----DYTAQONA-----KREKRTLF--DNU	1298
Qy	496	LHRGHEEDYKEHLLGASIYEAQLLKYYKKEKENALCSIIQNSYADLADIIGSDIIKDY	555
Db	1299	SHCEIQEYRDN-----KKYRWK--RSQTCs-----NWRKFSNQRTUYDY	1337
Qy	556	YGKQMEENLNKVNKKRNBESLKIIFREKWDENKENVWYMSAVLNKXKCTCKDYKDFQK	615
Db	1338	--RNWNNLKNKN--SYLRNLKSKIFLDKSVRKRKs-----SGTYKGSIDICKGIEK--	1386
Qy	616	IPQFLRFWEKWDGDFCEKREKIIYSPSFVKYCKKQCDENTC-----KNKSEYKWKIDL	671
Db	1387	-KETRW-----KRLHMKGSNNRNIIFSRIKDFKNSIIYNNQDKTNDY-----	1429
Qy	672	KKSEYEQV-----DKYTKDKQNKQWYDNIDEVKNKEANVYLKESKECKDVNPPDKIFNE	726
Db	1430	-XNGYSNNNNNNNNNTYNEKINSFN-----KRHSVAFKQMSKE-KLIPINNVLsN--	1480
Qy	727	SPNEYDMCKDEIKYLN--EIKYPTKWHIDYIDTDSFTFGDGTPIsINANINEQOSGK	785
Db	1481	-----RCDVESYPNEBSFSKKKKNSIvSPA-----SGK	1511
Qy	786	DTsNTGNSETSdSPVSHEPESDAALNVEKLSGDSSESSETRGILDINDPsvTNVNEVHDA	845
Db	1512	--SDIGFEERTDAGI-----AGKNMELSYNELEKYKYVKHIDSNNKYNNNNNSIH--	1560
Qy	846	SNTQGSvNTSDITNGH-----SESSLN-----RTTNAQDIKGR	880
Db	1561	---MSHSSNNNTFESHYIPDAIKYQSSLDDECLYSSSFLGVLVLNRKNTWCHEIDG	1617
Qy	881	SGNEQSDNQENSHSSDNQSLGTIGQVPSEdNTQYDSQNHPRDTPNALASLPsDDKIN	940
Db	1618	SPNKWTIIGVN--BFTNRGRQMSIvKXP-----DSM-----ESGSILVYKGSdSsIL	1662
Qy	941	EIEGFDSSR--DSENGRCdTTsNTHdVRt-----NIVSERRVNSHDfIRNGMANNNAHQ	994
Db	1663	SLLNLKYSKFLDKYsRRKRKNTQYKQKREYNElKsKDINTEEFQKSLS-----	1716
Qy	995	YITQIENN--GIIRGQ--EESAGsVN---YKDNPKRSNFSSENDHKCN-----IQEYN	1041
Db	1717	-LKNFKNTPKVLGECTLDNNNNNNINKNYKYDKNDKHNNNNNNNNNNsVSYNYLHDSN	1775
Qy	1042	SRDTRVRBEIILKsKQNKCNNEYSMBYCTYSDERNSSPQPCs--REERKKLCCQISDYC	1099
Db	1776	NLDAMQ-----NOKSYAVEDKEDYVYTNGBECNKYKERYRLEKQLRKFS	1821
Qy	1100	LKYFN--FYsIEYNYCISKEIKsPEYKCFKSEGOSSIPYAAAGILVIVILLSSsARMG	1157
Db	1822	VKGURSIFAPRYL-----SEETIKYRMWDACSSI-----YNKEQRLE	1862
Qy	1158	KSNEEYD-----IGESNTEATFEEN-----NY	1179
Db	1863	KVAEEFERDLIYLGITGVKNGLOEQKVPKPTIDILNQSGIRIWMLTGDNVVSLHVSFLCKP	1922
Qy	1180	LNLKsRIFNQEVQETN	1195
Db	1923	LNKHTKIFHAALENS	1938

[illegible]



```

; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvelle
; APPLICANT: No. US20040062769A Utaha. Fujii
; APPLICANT: Joseph Smith
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; TITLE OF INVENTION: (PFPMPI) THAT MEDIATES ADHESION TO CHONDROITIN SULFATE A
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: PCT/US00/24195
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/152,023
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; US-10-087-013-7

Query Match 5.4%; Score 348; DB 4; Length 294;
Best Local Similarity 32.9%; Pred. No. 1.4e-14;
Matches 94; Conservative 37; Mismatches 87; Indels 68; Gaps 15;

Qy 456 CNYKSNNEKSWNC--TGFTTNKFPQTCPPRRQTICLGRTYLLHGHEDYK-----EHL 508
Db 13 CNAKRRKNEWQCQDNFTVDNEGVCMPRRKRSICIHNLTL-----EQTKNKYQLREAF 67
Qy 509 LGASIIYEAQLLYKYKEDENALCSIIQ-----SYADLADIIGSDIIXDY 555
Db 68 IKCAAKETNLWDYK--NDKNEABELUKGKIPDFWMRIFPTGDFRDFCLENDMGKV 126
Qy 556 YGKMEENLNKV--NKKRRNEESLKIIFREKWDENKENVKMSAVL-----KNKE 605
Db 127 --DKVKKNIKVFNNSSKRGFK--KIDPENWNNENGPIWNGMLCALIHADTKDSIKND 182
Qy 606 TKRDYDK-----FQIPQLRWFKEGDDFCBKRKEKIIYSPESFKVEC 648
Db 183 NYK--YEKVTILAKRDGNGMTLSEFAKPKPLRFVWEYDDYCKEROKYLTEVAS---TC 238
Qy 649 KKKD-----CDENTCKNCKSEYKKWIDLKSEYKQVDKYTKDNK 689
Db 239 KSIDGGQLKDRG--CNNKDEYKIKMKKKKEEWNLQ--DKYIKQRE 282

RESULT 25
US-10-732-923-8668
; Sequence 8668, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8668
; LENGTH: 2719
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; US-10-732-923-8668

Query Match 5.2%; Score 338; DB 5; Length 2719;
Best Local Similarity 19.8%; Pred. No. 7.5e-10;

```

```

Matches 305; Conservative 199; Mismatches 490; Indels 548; Gaps 70;
Qy 33 NRQESSDISRVNPELNNHKTNIYDSY-----EDVNNKLINSFVENKSVKKRSLSFI 88
Db 1303 NKKGDDGDDILVSKNDEDDDDDDNSVDKYNNNNDNDAIDDEVMMDEAELESNNNT 1362
Qy 89 NNKTSYDIIPPSYSYRNKPFNSLENDNSGNTNNTNNTFANTSEISIGKDKO-----141
Db 1363 NNNNNNDGTGGGEMINGQFSS--SNNQNNNNNNNN-----DISLWRSVRKKPKQIL 1416
Qy 142 YT-----FIQKKT-----HLFAC-----154
Db 1417 YTPYGRQEVKQIKHEVRCFICNNTKHKALVKDRNGNDIDYGDMDINCFRCPTYH 1476
Qy 155 ----GIRKSIK--WIC-----RENSE--KITV-----CVPRKIQLCV 185
Db 1477 KLCEGIKDNVKTWTCWSWHECCCLCFKSSQCCNLIIHCATCPTSFYCNCFPDPYVRYV 1536
Qy 186 AN-----FINSRLETMEKFEIFLISVNTAKLLYNKNEGKDPIS 225
Db 1537 GEEYHNLRQGVNFTPNWVCFKSCKAVEBQK-----RRKMTKEEREHEKQ 1586
Qy 226 FCNELRNSFSDF-----RSSFI-----GDDMDFGENTDRVKGYNLK 262
Db 1587 LQKELRSQHLHDSKQEBLEAKKRAQOQLERKFIENRKRIDALDOQYEDQLKAYENVF 1646
Qy 263 FSDYKBNVKNLNNIKKEWKNKANLNMHMTVNHKGNISKECAIIPABEQINLWIK 322
Db 1647 PNNFVK-----ELVRIEHAQMLAQKIGVEDN-----NNNNNNKNTN-----KK 1686
Qy 323 WNFENFLEK---KRLFL--NIKDKVENKYEACFGCRLP-----CSSYTSFMKSK 370
Db 1687 KSTTFLHTKLPSKLLVLCENCKLPCHANYKYP--GKCCYPPBELDKSYNNTSFSQMR 1743
Qy 371 TOME-----VLTNLYKKKNSGVDKNLFNDLFCNNKNDLD--DPFKNEKEYDCLC 419
Db 1744 DGEKRVSEGNDSKCSHIFKDKVEDSNM-----KNITVIOGNKIMNSSPHKGGT 1797
Qy 420 DCRYTATIIFSLNGPAKNDVDIASQINVDNL--RGFCNYSNNKESWNC-----TGTF 473
Db 1798 KCEPNESGTTSPSNDGKKDLMTVTSTNNNNNNNNNNNNNNNNNNNNNNNNNNNNSSST 1857
Qy 474 NKFPCTGPPRRQTICLGRTYLLHGHEDY-----KEHLL--GASIIYEAQLLYKYK 525
Db 1858 NK-----LMSNKGIEYNKDNDEYNNNESKDKIMKGVGTTFLSDGSMRFP--KD 1904
Qy 526 KDNAL-----CSIIQNSYADLADI 546
Db 1905 SPENNENNNNNNNNNNNKDSQPNNDNTSKTKRFLRAVCSKCTVQ--LQKLAHFR 1962
Qy 547 KGSDDII-----KDYQK--KMEENLNKVKDKRNEESLKIIFREKWDENKENVKMSA 599
Db 1963 KHCDKLTEBEKKEYDEKREKLKELINLLNKKIK-----EETHEBEYKNMS--2008
Qy 600 VLKNETCKDY--DKFQKIQFLWFKEWGGDDFCBKR-----EKIYSPESFKVEC 648
Db 2009 VFKFSLYSNYQDKADDDILEDCIRSMKWDLIADKKKIIKKEKEKENKKIYS-----2060
Qy 649 KKKDCDENTCKNCKSEYKKWI---DLK-----KSEYKQVD-----KYTK 685
Db 2061 -KQDDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2119
Qy 686 DKNNKQMYNDIDEVKNKEANYLKEKSKCKVDNFDKIFN-----ESPNEYED 733
Db 2120 DMKEKINDELCTISGKRKKKIVKQIKKVKKKGGNDNENNNNNNNNNNNNNNNNNNNNN 2179
Qy 734 MCKKCD--EIKYLNKYP---KTKHDIYDIDTFSDTFGDG-----TPISINAN 777
Db 2180 SSKKDDDKLKLNLINPPHEKLIKDTNANFIRDSFNEDFLNKAKEFLRVTKSNLINN 2239
Qy 778 INEQQSGKQTSNTGNSSETSDSPVSHPEPESDAAINVEKLSGDESSETRGILDINDPSVTN 837
Db 2240 GNNNNNNKKSSNSK-----NINKNS-----SNTK 2266

```



```
Db 1419 -----ENDNNMNMENKGYNN-----SITVANKINKINH-----NFQ 1452
Qy 1076 RNSPGPCSRERKLLCCQISDY-----CLKYFNFYSEYNYNCIKSEIKSPYKCP-- 1126
Db 1453 NNIKTLSSEKKYMLTIDSINNSFNILDKVNTFHSYHIALPFLCKYTSYNFLH 1512
Qy 1127 ---KSEG-QSIPYFAGGTIV-----VIVLLSSASRSGKSNEEYDICESNIATPEE 1176
Db 1513 PINKNEHVSSILNSKNDILTDNGCNFVLYLLSFL-----SNKISSLRKCVANALY-- 1565
Qy 1177 NYVNLKLS-RIFNOEVOET-----NISDYSEYNYNEKNMY 1210
Db 1566 NSVLLQMPIRFRHNNLTKYGPNGYRIGYDIYKIANAFVNDYSTSEYKRDILY 1620

RESULT 27
US-10-087-013-9
; Sequence 9, Application US/10087013
; Publication No. US20040062769A1
; GENERAL INFORMATION:
; APPLICANT: Arthur Scherf
; APPLICANT: Louis H. Miller
; APPLICANT: Benoit Gamain
; APPLICANT: Dror I. Baruch
; APPLICANT: Pierre Buffet
; APPLICANT: Christine Scheidig
; APPLICANT: Jurg Gysin
; APPLICANT: Bruno Pouvellet
; APPLICANT: No. US20040062769A1utaka Fujii
; APPLICANT: Joseph Smith
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1
; FILE OF INVENTION: (PFEMP1) THAT MEDIATES ADHESION TO CHONDROITIN SULFATE A
; FILE REFERENCE: NIH176.001C1
; CURRENT APPLICATION NUMBER: US/10/087,013
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: PCT/US00/24195
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/152,023
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-013-9
```

```
Query Match 5.1%; Score 333.5; DB 4; Length 351;
Best Local Similarity 27.9%; Pred. No. 1.1e-10;
Matches 92; Conservative 46; Mismatches 131; Indels 61; Gaps 11;

Qy 442 IASQINVDLRGFCGNYKSNKSNCTGTFTNKFCGTCBPRTQCLG-RTVLLHRGH 500
Db 9 ISQIEKNH-----NCKTTEDAKWKCENTKLGDEGVCMPPRQNLGVHLYTKLNDSDK 64
Qy 501 EEDYKHLGASIEAQLKYKYEK---DENAL-----CSIQNSYADLADIKG 548
Db 65 EEDLRFAFIKSAAEFTLLRQYNSKNVEDDKILHRDIMPPEFRSMFYTFGVDYRDLCD 124
Qy 549 SDI---IKDY-----YKKMEENLNKVNKKRNEESLKIPIREKWDENKENVK----- 595
Db 125 TDISEKIADHVTAKKITTAVFOKIGSKTTNGKKVLE---REGWKEYGSLSIWKGMALCAL 182
Qy 596 -----VMSAVLNKQKTCQDYDFOKIPOFLRWPFKEGDDFCERKEKIY 639
Db 183 SYNTETKKMDGVRTYLMKYIKNDIKEVLEEFASRPPLRWTEWGEDFVNRKKELV 242
Qy 640 SFSEFKVEK-----KKKQCDEN-----TCNKKCSYKKWIDLKKSSEYKQVDKYTKDN 688
Db 243 SLKKKCDCTLRNNGTSNKTCDNENGCAKCTQCEKYKKMERWKGHYSQKKKFLQYKN 302
```

```
Qy 689 KMYNDIDEVONKEANYVLYLKEKSKCKDVN 718
Db 303 SATYNN--GLAVKEANSEYTKNDPEVTEAN 330

RESULT 28
US-10-732-923-3342
; Sequence 3342, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3342
; LENGTH: 2910
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-732-923-3342

Query Match 5.1%; Score 331.5; DB 5; Length 2910;
Best Local Similarity 18.5%; Pred. No. 1.9e-09;
Matches 289; Conservative 249; Mismatches 522; Indels 499; Gaps 70;

Qy 2 KGYFNIVFLIPLFLVNV-IRINESIIGRTLNYR-----QDESSD-----ISRVNSPELN 50
Db 425 KSISSNVNFKTLFLFVNVLLKVSVHLSG--LYTFMDILNYSDDLNNHNSNNNNNN 482
Qy 51 NNHKTNIYDSY-----EDVN--NKLINS-----FVE 75
Db 483 NNNNNNNNDYFISVELRKLIESINIYQAVNAALYLFHYILHMLEITSQKNNMLIE 542
Qy 76 NKSVKKRSLSFINNTKSYD--IIPSYSYRNDKNSISEDNDSNTNNTNNTSEI 133
Db 543 NE-ISTIRNFIKINKKVKVDSNILINIVQYKN-----YPNTQNNNTY----- 585
Qy 134 SIGKDNKQVTFIQKTHLFAAGIKRKSIKWICENSEKITVCVDPDKIOLCVANFLNRL 193
Db 586 ---KSNKQLNIQNVENKKYKYNLKE---LVHRNFE-----IRKILICKY--NKII 630
Qy 194 ETMEKPKEIFLISV-----NTEAK---LLYKNVEGKDPs----- 224
Db 631 KWMKKIHLNYIMHVYKSYKNDQNGYINBEKNETKTETKNEKQVETKREINNEKNETKR 690
Qy 225 IFCNELRNSFSDFRSSFIGDDMDPGGNTDR----- 254
Db 691 EINNETKKERDNRDNRDNRNERNERNERNERNERNERNERNERNERNERNERNERNERN 750
Qy 255 -----VKGYINTKFSDYKKNVEKLN-----NIKKEWBEQKANKLWNHMIVNH 298
Db 751 YQTPYCCNITYTKDKIENKKNYNIPEVESLSLFFHCCNINTYTYKKEIDINMKLIYF 810
Qy 299 KGNISKECAIIIPAEPOQINLWIKEMNENFLMEKRLFLNKKD-----KCYE-----NKKY 348
Db 811 LSNLL-----TKYQNSEFIKROMNSYFKEINNLFDFDKIDYDILLCYNNMRNLNKE 863
Qy 349 EACFGGCRPLPCSSYTSFMKSKTKTOMEVLTN-----LYKKKNSGVDKNNFLNDFKKNKN 403
Db 864 NINIPTCOL--NNSTCOLANNSTCELANNSTNQLNFTYNKDLCSKEIEDLLSDMSKYNIN 921
Qy 404 DLDDFFKNEKEYD-----DLCDCTATATIKSFLNGPAKNDVDVIA----- 444
Db 922 -VDKIMHKYNTYNDIELLKSNMKNVDMLLLLSNNIYNKIMSTTYLNNFDILKFMKI 980
Qy 445 -----QINVNDLRGFCGNYKSNKSNCTGTFTNKFCGTCBPRTQCLGRTYLLHRG 499
Db 981 FVHVIKIFVKDYTML---YINNQGKINESNEDHN-----LNKNKDSHID 1022
Qy 500 HBEDYKEHLLGASIEAQLKYK--YKEDENALCSIQNSYADL--ADIKSGDIIDKY 555
```

Db 1023 NNIENFLNQLIDIEVEKRVNIIKVELNEKNIIINNIKELOEKEIKYNDIKTNL 1082  
Qy 556 -YKOMEENLN-KVNKKKNEBSLKFREK---WWDENKE-----591  
Db 1083 EREKKNMHNIELEKREKINDENMIKFEKKNKMYDINDKELENEKKNKEIYWELONIK 1142  
Qy 592 NVKVMASVLNKK-----ETCQDYDFQKIQPOFLRMFKWGGDFCEKREKEIYSPFSKV 646  
Db 1143 NINSDMNPALDEKSVQDQKCIDPEKLNKM-----YEDTCKELNNKNVLYEQMNT 1192  
Qy 647 ECKKDCDENTCKNKSSEYKWDLKK-----EYEQVDKVTKD-----686  
Db 1193 QLYKEKINNDI---CTELEKIKNINKNQIINLENEKKTNEQKQIETKOKMKIKEMVIQ 1249  
Qy 687 -KNKKMYDN--IDEVKNKEAN---VYLKEKCKDQVDFDKIFNESPNEVECMCKCDE 740  
Db 1250 LEKEKINNEIITOLENEKLNKIY-----EDIEKKNKLNSELENYEQKKINE 1301  
Qy 741 I-----KYLNE---IKYPTKTHDIYDITFSDTGDGTPPISANINPEQSGKDTST 790  
Db 1302 MIQLEKEKIINNEIITOLENEKSVKINTELENIK-----KMDENMEKLNKEQKINN 1356  
Qy 791 G-----NSTSPSVHSEPEAAINVE-KLSDSESSSETRI-----827  
Db 1357 DLQDLENEKNEKDHINKEFEKIKNIQIKQNMDEKKTNEISLNVIEDRLYKINEHVLL 1416  
Qy 828 -----LDINDPSVTNNVNEVHDASNTQGSVNTSDI-----858  
Db 1417 NLEBEKKIDQFTIDLENERKRNHILQHPDDEKKNQNDQYEDLLKEREKKNKLQILIYDD 1476  
Qy 859 -TNGHSESL---NRTTNAQDIKI---GRSGNEQ-SDNOENSHSDNSGSLITIGQVPS 909  
Db 1477 GKNNISQNLIDLENERTRNDLAILDQEKKQNEQINDLENERKNNQONLQNLNEQKK 1536  
Qy 910 EDNTQNTYSQ-NPHRDTNALASLPDDKINIEGDFDSRDSSENGKGTTSNTHDVRT 968  
Db 1537 KEQNAVSYEBQKINHOLENELOKQRTYK-KMAKPERKFLMKN-----TNDTQIKOT 1590  
Qy 969 N-----IVSERRVNSHDFI---RGMANNNAHQYITQI-----ENNGIIRQGESAG 1013  
Db 1591 QQIITQKIITQIKITQIKITQIKITQIKITQIKITQIKITQIKITQIKITQIKIT 1649  
Qy 1014 NSVNY-----KONPKRSNFSENHKKNI-----QEYNSRDTKRVREBI-----1052  
Db 1650 KQVNDMIQIKQKRIDELNLEKEKEVNDKIIIQEYQMK-JEHINELEKEKEI 1708  
Qy 1053 -----IKLSQKNNKNNSEYCTYSNDRNSPSPGCSREERKLLCC 1093  
Db 1709 NNNLIRIEKNEKNEQNLQIDTEKKNQMSIE-LDEKKNQHRRTELDEKKNH 1766  
Qy 1094 QISDYCLKYNFYSIHYNCIKSEIKSPYKCFKSGQSSIPYFAAGGILVIVLLSSA 1153  
Db 1767 EMN-----IBLDEEKNKE-----INILLEENK 1790  
Qy 1154 SRMGKN-----EYDICE-SNIRATPEENLYNKLRIFNQYEQETNISDYSVYNEK 1207  
Db 1791 NKIIQINSLCKEKEINELLNIQLEKKN-----IQELNIKLEDEKKNQ 1837

## RESULT 29

US-10-732-923-22709  
; Sequence 22709, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149

; SEQ ID NO 22709  
; LENGTH: 3124  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-10-732-923-22709  
  
Query Match 5.1%; Score 329; DB 5; Length 3124;  
Best Local Similarity 20.1%; Pred. No. 2.8e-09;  
Matches 246; Conservative 20; Mismatches 434; Indels 344; Gaps 60;  
  
Qy 2 KGYFNIFLPLIFLVNVRINESIIGRTLYNNRQDESSDISRVNSPELNNHNTNYDSD 61  
Db 1736 KGRMKTLYLLVDI--LDVVRKGESL-----NYSSS 1764  
  
Qy 62 ---YEDVNNKLINSFVENKSVKKRSLSPFNNTKSYDIIIPPSVYRNDKFNLSNEDN 118  
Db 1765 NLLLSQLGSEANSIYEREDI-KEGSDI--KESRDII-----KESRDIIKEISTN 1815  
  
Qy 119 SGNVSNNTFANTSEISIGKNDKQYTFIQRTHLFACGIRKRSI-----KWICHENSEKIT 173  
Db 1816 ISKSSSRNISKSSRSI-SDIKEGQIIDKEDLIFKINRMKNKIDSRYSKRIDKESRDKI- 1873  
  
Qy 174 VCVPRDKIQLCVANFLNSLETMEKPEIFLISVNTAEKLLYNKEGKDPISFCNELRNS 233  
Db 1874 -----SDKTNHVLDEVVHSDIHLN-----YEINNR-----CKMKG 1908  
  
Qy 234 FSDPRSSFTGDDMDFCGNTDRVKGYINTKFSYDYKKNVEKLNKNIKKEWKNKANLWNH 293  
Db 1909 TNN-ENKLGIDIFNM--YDKIKYI---YKQYKSKSMENISFICH--YRNTKYKSDY 1959  
  
Qy 294 MIVNHKG-----NI-----SKECAIPABE---PQINLW 319  
Db 1960 LLLDNKESKPKRNTSYVLESPLHLIGDIVDNNIKRKKKKEIKTIVSDMFTSPVN-- 2017  
  
Qy 320 IKEMNENFLMEKKRLFLNKKOVENKCYEACPGGRLPCSSYTSFMKSKTKTOMEVLTNL 379  
Db 2018 IKEYNTN--EOER-----KGEIVGNLSYDKTKIC-----PPIKFTKEGR---IKGNK 2060  
  
Qy 380 YKKNSGVDKNNFL-NDLFKK-----NNKNDLDDFFKNEKEYDLDCCRVTATIKS 430  
Db 2061 IEKKEKCYNNFLYNDYSSSPKYGDNENNFKYIRERKDPKQKFD--HPNENFSK 2118  
  
Qy 431 FLN--GPAKNVDVDAQINVDLGRFGCNYKSNKESWNCCTGFTNKFPGTCPPRRQTL 488  
Db 2119 FLHNTYPMNK-NKNNKNNVNRNEYPNYTSKDGVSYNFLSDLSLSSDNE-----2170  
  
Qy 489 CLGRTYLLHRGHEEDYKHLGLGASIIYEAOLLYKKEKDENALCSII-----QNSYA 540  
Db 2171 -----YSSDNEYSDDSEKY-----YKRVKKNKIIKFDLFTKIYIEKRLQNMNYD 2219  
  
Qy 541 DLADIIIGSDIIDKYGKKEENLNKVNKKRNEBSLAFREKWDENKENVKUMSAV 600  
Db 2220 -----YKGR-----GKLLKNGKMERNTKYKNVN--BITKMKYFVNNENRDHEV---2261  
  
Qy 601 LKNKE--TCQDYDKFQKIQPOFLRMFKWGGDFCEKREKEIYSPFSKVBC-----KKDCDE 655  
Db 2262 --NKEDISKMQY-----FLHSHKKEQIEDKXTHYFKN--VECVYPAGNNINH 2312  
  
Qy 656 NTCNKKCSYKWIIDLAKSEYEQVDKYTKDNKKNYDNIDEVNKEANYLEKSKCK 715  
Db 2313 NPSREKRYK--INL-----YD-HLDEQEKIKGKKYFNKDKELIGSINKQTERKPKKN 2365  
  
Qy 716 DVNFDKIFNESPEYEDMCKDEIKYLNELIKYPTKTHDIYDITFSDTGDGTISIN 775  
Db 2366 KKNENK-----KOKKIRMTNNTKKEKSHSIIISVE-----EQN 2400  
  
Qy 776 ANINEQSGKTSNTGSETSDSPVSHPESDAAINVEKLSGSDSESSETRGILDINDPSV 835  
Db 2401 MHNNSLKKKEVNFQKNE-----EYLNANTNCSLIGKEWEEDVYEFHNNIYNNQTSY 2456  
  
Qy 836 TNNVNEHDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSGNEQSDNQSHS 895  
Db 2457 SDDIN---NTTKLKGNNNTNDISKNGKNLKG-----KISFFSMNNKYHESEIMNE 2506

Qy 896 SDNSGLTIGQVPSDNTQNTYDSONPHRTPTNALASIPSDDK----- 938  
Db 2507 EDNKNMNLNLTQ-----SQIINKDKYNYFTCHPSLKKKCKSVFTKINNLFKNYFK 2554  
Qy 939 --INIEGFGDS-----SRDSNGRG-----DITGNTHTDVRRTNIVSERRVNSHD 980  
Db 2555 SIDVHEKFGSKFKFHSKSDDIKGNKKISKRYNNNNNNNSYNSIDSGK--YSHN 2612  
Qy 981 FINGMANNNAHQYITQIENNGIIRQBSAGNSVNYKONPKRSPFSSDHHKNIQY 1040  
Db 2613 NKKNHHNNKKYHH-----NN-----NKVHHNNKKYHH--QNNYKKGHHN 2654  
Qy 1041 NSRDTKVRREILKSKONCKNNEYSMEYCTYSDERNSSFGPCSRERKKLC--CQISDY 1098  
Db 2655 NSR-----VMSLGEKTEKXENVDAIYQFDNYD-----KKLLKKLTSLNQLNKK 2698  
Qy 1099 CLKYFNFSYIEYCNKISEIKSPY 1123  
Db 2699 NVKNFNF-----YKFNDELEEE-EY 2719

RESULT 30  
US-10-087-013-11  
; Sequence 11, Application US/10087013  
; Publication No. US20040062769A1  
; GENERAL INFORMATION:  
; APPLICANT: Arthur Scherf  
; APPLICANT: Louis H. Miller  
; APPLICANT: Benoit Gamain  
; APPLICANT: Dror I. Baruch  
; APPLICANT: Pierre Buffet  
; APPLICANT: Christine Scheidig  
; APPLICANT: Jurg Gysin  
; APPLICANT: Bruno Pouvelle  
; APPLICANT: No. US20040062769A1utaka Fujii  
; APPLICANT: Joseph Smith  
; TITLE OF INVENTION: IDENTIFICATION OF THE DOMAIN OF  
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ERYTHROCYTE MEMBRANE PROTEIN 1  
; TITLE OF INVENTION: (PEMPTI) THAT MEDIATES ADHESION TO CHONDROITIN SULFATE A  
; FILE REFERENCE: NIH176.001C1  
; CURRENT APPLICATION NUMBER: US/10/087,013  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: PCT/US00/24195  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/152,023  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 308  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum  
US-10-087-013-11

Query Match 5.1%; Score 327.5; DB 4; Length 308;  
Best Local Similarity 27.8%; Pred. No. 2e-10;  
Matches 85; Conservative 46; Mismatches 118; Indels 57; Gaps 10;  
Qy 455 GCNPKSNKSNWCTGFTWKFGCTCEPPRPTILCLGRTYLLHRGHE-----EDYKEHLG 510  
Db 10 GCNPK-ESYDPDCKGNIDSHSGACMPRRQRLCVRD---LTQGGIRKPEIDILTKFIN 65  
Qy 511 ASIYEAQLLYKYKEKDENALCS1-----IQNSYADLADIKGSIDIIKDYGK 558  
Db 66 CAAKETHFAHMYKKNVNAENELKSGKIPGFRKQMYTFGDRDIFFGTDISSCRYIK 125  
Qy 559 KMEENLVNKKVKKQNE-----ESLKI PREKWDENKENVKVMGSAVLKN-----K 604  
Db 126 DTSQTIKSLGDAQATTEKGDTHDDNKKL--QEWWTTHGPKIWEGLMCAITNGLSSEKK 183  
Qy 605 ETCKDY-----DKFOKIPQFLRWFKWGGDDFCRKRKEKIYSFESPKVECK 649

Db 184 NILQDYSYNKLNAEKDCCLCFASKFQFLRWYVWVSDEFCECRKKLEKVEDVICAK 243  
Qy 650 -----KKDCDENTCKNKCSEYKKWIDLKKSEYEKQVDKYTKDKKKKKYDNIDVQNKAN 704  
Db 244 DYEGCKNNKSNNSCVKCKEYENYITGKTKQYBSQSGKFKTEKQKK-PEYNSYSKKDAS 302  
Qy 705 VYLKEK 710  
Db 303 EYLKOK 308

RESULT 31  
US-10-732-923-22588  
; Sequence 22588, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 22588  
; LENGTH: 3127  
; TYPE: PRT  
; ORGANISM: Plasmodium falciparum 3D7  
US-10-732-923-22588

Query Match 5.0%; Score 326.5; DB 5; Length 3127;  
Best Local Similarity 19.9%; Pred. No. 3.9e-09;  
Matches 246; Conservative 198; Mismatches 428; Indels 365; Gaps 61;  
Qy 2 KGYFNIELPLIFLYNVIRINESIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDSD 61  
Db 1736 KGKMTYLLVDI--LDDVVRKGESL-----NYSSS 1764

Qy 62 ---YEDVNNKLINSFVENKSVKKRSLSPINNKTYSYDIIPPSYSYRNDKFNLSNEDN 118  
Db 1765 NLLLSQLGSEAVSYIEREDI-KEGSDI--KESSRDI--KEDSRDIKEISTN 1815

Qy 119 SGNNTSNFANTSEISIGKDNQYTFIQKRTHLFACGIKRSI-----KWICRESEKIT 173  
Db 1816 ISKSSRNISKSSRSI-SDIKEGQIIDKEDLIFKINRMKNKIDRSYKIDKESRDKI- 1873

Qy 174 VCPDRKIQLCVANFANLSLETMEKPKFIPLISVNTAEKLLYNKNEGKDSIFCNELNS 233  
Db 1874 -----SDKTNHVLDEVVRKSHDIHLN-----YEINNK-----CKMKGD 1908

Qy 234 FSDFRSFIGDDMDFGNTDRVKGYINTKFSYDYKEKNVEKLNKIKKWEKKNKANLWNH 293  
Db 1909 TNN-ENKLGIDIFNM--YDKKIKYI--YKKNYKSKSMENISFIKH--YRNTKYKSDY 1959

Qy 294 MIVNHKG-----NI-----SKCALIPABE-----PQINLW 319  
Db 1960 LLLDNKGSKKFKRNTSYVLESPLHLIGDIVDNNIKRKKKKKEIKTIVSDDMFTSPVN-- 2017

Qy 320 IKEMNENFLMEKKRLFLNIK-DKCVENKKEACFGCRLPCSSSYTSFMKSKTQMEVLTN 378  
Db 2018 IKEYNTNEQERKKEIVGNLSYDK--TKKI-----PFFIKTKEGRKKKKIEKKEK 2066

Qy 379 LYKKNKSGVDKNNFL--NDLPFK-----NNKNDLDDFFKNEKEYDDLCDCRYTATIIK 429  
Db 2067 KEKKENN---NNFLYNDYDSSYSSPKYGDNNENFVIKIRERKDFQKKFD--HPNPNFS 2120

Qy 430 SFNL--GPAKNDVDIASQINVNDLGRFGCMYKSNNEKSWNCTGTFNTKFGCTCEPPRQT 487  
Db 2121 KFLHYNPMKN-----KKNKK--NNKNVRRENEYNYTSSSKD-- 2156

Qy 488 LCLGRYTLHRGHEEDYKEHLLGASIEAQLLYKYKEKDENALCS--IQNSYADLADI 545  
Db 2157 --GVSY-----NFLSDSLFSSD---NEYSSDNEYSSDSEKYYKKFKKNKKI 2198

Qy 546 IKSDDIIKDY- - - - - GKMEENLKNVNDKRNESLKIPIREKQWDE 588  
Db 2199 IKPDDLFTKIYIKKRLLOMNYDVKGKGLKNKGNERNKTKYKNV- - - - - EITMKYFVN 2256  
Qy 589 NKENVWMSAVLQKKE- - - - - TCHDYDKFOKIPOLRWFKEWGGDFCEKKEKIKYISPEFKVE 647  
Db 2257 NENRDHEV- - - - - NKEDISKSMQY- - - - - FLHISKHKEQIEDKCKTKHYFHN- - - - - VE 2303  
Qy 648 C- - - - - KKKDCDENTCKNKSCKYKWKIDLKKSEYKQVDKYTKDNKKWYDNIDEVKNKEA 703  
Db 2304 CVYPAGNINHNFSRNEKRIS- - - - - INL- - - - - YD-HLDEQEKIKGKKYFNKOKELIGSI 2356  
Qy 704 NVYLKESKECKOVNPDQKIFNESPEYEDMCKDEIKYLNBIKYPKTKHDYIDIDTFS 763  
Db 2357 NKQTERPKKKKNKNIENK- - - - - KKKKIRMITNKTKEKHSNLSIISVE- - - - - 2400  
Qy 764 DTFDGTGISINANINQOQSOTNTGNSSTSPVSHPESDAAINVEKLSGDESSE 823  
Db 2401 - - - - - EQNMHNNSLKKKEVNFCKNE- - - - - EYLNFRANTNCSLGIKEMEEDVYEFH 2447  
Qy 824 TRGILDINDPSVTNNVEHVDASNTQGSVNTSDITNGHSESLNRTTNAQDIKIGSGN 883  
Db 2448 SNNIYNNQTSYSDIN- - - - - NTTKLKGMGNNTDISXNKGKGLGK- - - - - KISPFM 2497  
Qy 884 EQSDQENSHSSDNGSLTIGQVPESDNTQNTYDSQNPHRDTPNALASLPDDK- - - - - 938  
Db 2498 NKKYHSEIEMEEDKNWLNLTQ- - - - - SQIINKDKNYFTHCPSLKXKXSV 2545  
Qy 939 - - - - - INEIEGDS- - - - - SRDSNGRG- - - - - DTTSTNTHDVRT 968  
Db 2546 TKINLNFKNVFKSIDVHEKFGSKFKFHSKSDDIKGNKKNKISKNYNNNNNNNNYS 2605  
Qy 969 NIVSERVNSHDFTRNGMANNNAHOVITQIENGIIIRGQESAGSNVYKDNKPRNFS 1028  
Db 2606 NIDSGK- - - - - YSHNNKGNHNNNNKYYH- - - - - NN- - - - - NKYHHNNNNKYYH- - - - - 2645  
Qy 1029 SENDHKNIQYNSRDKTRVREYIILKSKQNKCNBYSMCYTSDERNSSPGCSEER 1088  
Db 2646 QNNYKHHNSNR- - - - - VMLSKGKTEKNVYAYQFDNYD- - - - - KKL 2689  
Qy 1089 KLC- - - - - QOISDYLKYNFYSIEYNCISIKSPY 1123  
Db 2690 KKLTSNLQNLKKNVKNFMF- - - - - YKFNDELEB-EY 2722

## RESULT 32

US-10-153-273-10  
; Sequence 10, Application US/10153273  
; Publication No. US20020169305A1  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; Chitnis, Chetan  
; Miller, Louis H.  
; Peterson, David S.  
; Su, Xin-zhaun  
; Wellens, Thomas B.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/153.273  
FILING DATE: 21-May-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/210.288  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Fuller, Michael  
REGISTRATION NUMBER: 36,516  
REFERENCE/DOCKET NUMBER: NIH121.1FWDV1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 700 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmodium falciparum  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-10-153-273-10

Query Match 5.0%; Score 325; DB 4; Length 700;

Best Local Similarity 24.4%; Pred. No. 7.6e-10; Indels 140; Gaps 22;  
Matches 120; Conservative 70; Mismatches 161;

Qy 478 GTCEPPRRQTL- - - - - LGRTYLLHRGHEEDYKEHLGASIEAQLLKVKY- - - - - KE 525  
Db 8 GACAPYERLHLDYNLESIDTTSTTHKLLLE- - - - - VCMAYEGSINTHYTQHTNE 62  
Qy 526 KDNALCSIIQNSVADLADIKGSDDIKY- - - - - YGKQMEENL- - - - - NKVNDKRRNE 575  
Db 63 DSASQLCTVLARSFADIGDIVRGKDLVLYGVDNKEQKLEQKLDIFKKIHKDVMTN 122  
Qy 576 ESLKIP- - - - - REKWDENKENYWK- - - - - VMSAVLKNKET- - - - - 606  
Db 123 GAQERYIDDAKGGDFPOLREDWNTSRETVMKALICHAPKEANYFIKTACNVGKTNGQC 182  
Qy 607 -CKDYD- - - - - KFKIPQFLRFKEWGGDFCEKKEKIKYISPEFKVECK- - - - - K 651  
Db 183 HCIGGDVPTFYDVPQVLRWFEEWAEDFCRKKKGL- - - - - ENLQKQCRDYEQNLYCSNGY 239  
Qy 652 DCDE- - - - - NTKNKKSEYKKWIDLKSEYKQVDKYTKD- - - - - KNK 689  
Db 240 DCTKTIYKGLKLVIGEHCTNCSVMCRMYETWIDNQKKEFLKQKRYETISGGSGKSPK 299  
Qy 690 KMYDNIDEVKNKEANYLYKESKECKDVNFD- - - - - KIFNESPEVEDMC- - - - - K 736  
Db 300 RTKRAARSSSSDDNGYESFPYKGLKEVGQDVDFLKLNLK- - - - - EGICQKQPVQVNE 354  
Qy 737 KCDEIKYLNBIKYPKT- - - - - KHDYIDIDTFSDFDGTGISINANINQOQSOTNTGNS 794  
Db 355 KADNVDTNE-KYKTFSTRTEICEPCWCGLEKGGPPKV- - - - - KGDKTCSAKTK 405  
Qy 795 T- - - - - SDSPVSHPESDAAINVEKLSGDESSETRGIL- - - - - DINDPSVTNNVNE 841  
Db 406 TYDPKNITDIPVLPDKSQQNLKYNKFCCKGAPGGGQIKKQVCYYDEHRPSSKNNNC 465  
Qy 842 VHDASN- - - - - TQ 850  
Db 466 VEGTWKFTQ 476

## RESULT 33

US-10-732-923-3353  
; Sequence 3353, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D



```

; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-013-10

Query Match          4.9%; Score 319.5; DB 4; Length 311;
Best Local Similarity 30.0%; Pred. No. 5.7e-10;
Matches 94; Conservative 42; Mismatches 112; Indels 65; Gaps 15;

Qy 455 GCN--YKSNBK--SWNCTGTFNKFPGCTCPTPPRRQTLCLGRTYLLHRGHEEDY--KEHLL 509
Db 7 GCNKYKAGDKYPCWDCNSQIHITHNGACMPPRQKLCVSLGTLTDRKAIEVIRTEFI 66

Qy 510 GASIVYEAQLLYKKEKDNALCSI-----IONSADLADIIGSDIHKDY-- 556
Db 67 KSAAIETHFADRYKEDNGEAEALKNGNIPEGFKRQMYTFGVDYRDIFFGRDISTHAYI 126

Qy 557 ---GKMEENLNKYNKDK---KRNEESLKIFREKWDENKENVKWSAVL-----K 602
Db 127 SGVSPKVITILEKENDAKYAAKQNSNELL--DDWDOHGKDIWEGMLCALTHKISDEEK 184

Qy 603 NKETCKDY-----DKFOKIPOLRWFKEWGDDFCBCKRKEKIYSPESPKVE 647
Db 185 KKEIKNKYSYKLNESPKGSKNVKVEDFAKQPLRWFIEWGDEFCQAOREK---BAKVYS 241

Qy 648 CK-KKCD--EWT-----CKNKCSEYKWIIDLKXSEYKQVDKYTKDK--NKGMYNDIDE 697
Db 242 CSDAKYDGCNTKSNASCVSACKVYEDYTKKVEYTKQKFKDAEKITDKEGYEGF-- 299

Qy 698 VKNEANVYLKEX 710
Db 300 -STKDASEYLKCK 311

RESULT 35
US-10-732-923-3340
; Sequence 3340, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3340
; LENGTH: 1939
; TYPE: PRT
; ORGANISM: Plasmodium chabaudi
US-10-732-923-3340

Query Match          4.9%; Score 318.5; DB 5; Length 1939;
Best Local Similarity 19.9%; Pred. No. 6.1e-09;
Matches 282; Conservative 216; Mismatches 523; Indels 393; Gaps 68;

Qy 18 NVIRINSIIIG----RTLYNRQDESSDISRVNSPELNHNKHTNIDS---DYED---- 64
Db 208 NLEKINKVIEKENNRLKELKFMKEKNEI--IESLDGTINDKKNAYEKLKLEISFBEKRWI 265

Qy 65 --VANKLI---NSFVKNKSVKKGKSLSPFNKTKSYDIIPPSYVRNDKFNLSLENEDNS 119
Db 266 EMLDSKLIEKENFA-NKQAKLEKENEIIEKLKDIESREKDFKSKKEKFSAM--ENELNT 323

Qy 120 GNTN-SNNFANTSEISGKONQVTFQKTHLP-----ACGKRRKSI- 161
Db 324 LKSDLSKNACQMEYVKLEIKDLSQSLVEKEREIFEIKNEYDDKINNMMKELSLINDKGD 383

162 KWICRENSKITVCVPRKIQI-----CVANFLNSRLTMEKFEIFLISVNT 209
Db 384 NTVLHSEEEKINKLLKEKETELNEIHKYKNLEIETIKNELNEKEEELKNAKHTVEYN 443
Qy 210 ---EAKLLYNK---NEGKOPSPFCNBLRNSFSFRSFIIGDDMDFGNTDRVKG---YI 259
Db 444 LTKELIKLKEKTEDAKEGHKNEL--NELNQLSKLNK-----EKDNINKNTEPTEL 490
Qy 260 NTKFSDYVYKENV-----EKLNNIKKEW--WEKNKANLWHEMIVNHNKNI 302
Db 491 NDKLSSLNSEVNIINKQKQTLGNDIKYLDLNNLKNEINTSDNRNKNOKKEDLAWLNEEM 550
Qy 303 SKECAIPABEPQINLMIKEMNENFLMEKKRLFLNIKDKCVENKKYKACFGGCRLPSSY 362
Db 551 EGKCVVIDEIK-----KYKNEIFMLEEK-----LAEK-----ENY 581
Qy 363 TSFMKKSQTOMEVLTN-LYKQNSGVDKNNFLN---DLFKQ-----NKNDLDDF--- 408
Db 582 ADL-----NDEISILRNSIYVKEKEFIEMKPEYENKINLFNKNFEEKKNIYENELNSLRK 637
Qy 409 PKNE---KEVDDL-----CDCHYTATI---IKSFLNGPAKNDVDIASQINVDLRGFG 455
Db 638 YDNEQGLIKQIDELNIQIKLTKTEEKYLOLYNDNMHMFPSICTKIDMPYSENIGSDLVDFV 697
Qy 456 CNY-KSNNEKSWNCTGTFTNK-----FPGTCEPPRRQTLCL 490
Db 698 TAYIKRDESSDANPDTTHKEMVAELEKRAHAIVAELEKHEKEIAKLGSHKEVVLRL 757
Qy 491 G-----RTYLLHRGHE-----EDYKELLGASIEBAQL---LKYYKE-----KDNA 530
Db 758 GEQHKBTIILESKHVDVTKLGEQHKENIILKEEHKDVVTKLGDQVYKKEIAKLKEEHA 817
Qy 531 LC-----SIQNSYADLADIIGSDIHKDYKQWER--NLNKNVKNKKN 574
Db 818 VVVAELEKHKLGEGHKEMVDELEKRAHDFVEGLEKHEKHAETAKLGEGHSEMMNEVEKRH 877
Qy 575 EESLKIFREKWWDENK---ENVVMKMSAV-LKNKETCKDYDKFQK--IPQFLRWFKEWGD 628
Db 878 ADFVEGLEKHEKHAETAKLGEGHREVVALEKHEKHEVALEKHEKHEIAKLEEGHKEVMA 937
Qy 629 DFCRKEKIKIYSPESFKVECKKDCDNTCKNCKSEYKWIIDL-----KKSE 675
Db 938 ELGEKHEKHEVAGLEA-----KHNLEEGHKEMVALEKHEKRAHDLVAVLEEQHAKSIIKGE 991
Qy 576 YEKOV-----DKYTKDNKKNMYNDIDEV-----KNKEANVYLKESKECKVDNPDKIFN 725
Db 992 EHKEVVAGIIEKYKVEAKLAEBHKVYTKLGEQHKHEIAKLEGDGHE-----VYN 1042
Qy 726 ESPNEYEDMCKCKDEIKYLNIEIKYPTKTHDIYDITFSDTFGDGTPIISINANINEQOSGK 785
Db 1043 EVEKKNASLLNMLEE--NHKNEMIKLEKHEKESASDLVEKLY-----QKDE 1086
Qy 786 DTGNTGN-----SETSDS-----PVSHEPEPSDAAINVKLGSDSESSETRG 826
Db 1087 EVKNSNKKIEBLTNVTKDLNDSIMCYKKQILEEVEKREYNEERINKLQVNE----- 1139
Qy 827 ILDINDPSVTNNVNEVHDASN--TQGSVNSTSDITNGHSESSLNRTTNAQDIKLGSGNE 884
Db 1140 MKOMNDKXILKEKENEIKKLNKLSNYKVFETKENTYKNSVMVNE--NKEKIIIV----- 1191
Qy 885 QSDNQENSSHSS--DNSG---SLTIGQVPSEDNTQNTYDSQNPHRDTPNALASLPDSDKIN 940
Db 1192 DSVCKENISSEVGEKGNLKWTLKCKKERNIPFSINDKNSESELVDTIKSA-----YIN 1247
Qy 941 EIEGFDSSRSSENGRGDTTNTHDVRRNTNIVSERRVNSHDFIRNGMANNNAHHYITQIE 1000
Db 1248 KIEMKKKEIB-DNGK-----NIEDLK-----NKILDLSNELINLENKMKVLTDEN 1291
Qy 1001 NNGIIRQESAGNSVNYKONPKRNSFSENHDH-----KKNIQEYNSRDTKVEVEIILKS 1056
Db 1292 NN--LKKEBIKONKLNKEKENTEILNLDNDIILKLGSEIWEKDEEELKTKENIKLN 1349
Qy 1057 KONKCNNEYSWEYCTYSDERNSSPGCSREERKKLCQISDYCLKYFNFYFIEYVNCIKS 1116

```

```
Db 1350 DIEQINKEYI-----:|||:|-----KEENLM-----KEN-ENINEVTSKLN 1381
Qy 1117 EIKSPYKCFKSEQSSIPYFAGGILVIVLLSSASRMGKSNEEYDIGESNI-EATFE 1175
Db 1382 QIEIKMKLEELNKVE-----:|||:|-----LILAEKRETNMSISNDNKIVENNILEDTDS 1429
Qy 1176 ENNYLNKLSIFNOEVOETWISDYSEYNN-EKN 1208
Db 1430 KQNNLNK-----:|||:|-----NVEDKTGDDINCEKN 1451

RESULT 36
US-10-732-923-8666
; Sequence 8666, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8666
; LENGTH: 2110
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-732-923-8666

Query Match 4.9%; Score 318; DB 5; Length 2110;
Best Local Similarity 19.2%; Pred. No. 7.3e-09;
Matches 272; Conservative 185; Mismatches 468; Indels 492; Gaps 63;

Qy 33 NRODESSDGRSVSPENNNHKTNIYDSYEDVNNKLNSFVENKSVKKRSLSFINKT 92
Db 338 NNDEKSGDSIGEDDDNNHKGSDNNIDEDNN-----DHKSDNIGEDDDNNDE 390
Qy 93 KSYDIIIPSYRNDKFNLSLSEEDNSGNTSNNFANTSBSISCKDNKQYTFIQKTHLF 152
Db 391 KGGD-----NNIDENDN-NSDHKSDNIDENDNNSHQ-----DQEQFHETKDI-- 437
Qy 153 ACGIKRKSIKWICRENEKTIQVCPDRKIQLCVANFLNSLETWKEKEIFLISVNTAK 212
Db 438 ---IKNSSYEH-----DNK-----NYYN-----KTGEDYK----- 460
Qy 213 LLYNKGKDPISIFCNELNSFSDFRSFIDGDDMDFGNTDRVKGYINTKPSDYKEKV 272
Db 461 ---SDKENYSPTRFNKLKE-----:|||:|-----KYDEYDTKLKI 489
Qy 273 EKLNNIKKEWKNKANLWNNHMIVNHKGNISKECALIPABEPQINLWIKENENFLMEKK 332
Db 490 EKREENKN-YEKDEHEY-----ESDNYDKE-----KIN-----KKX 520
Qy 333 RLFL-----NIKDKCVENKYEACFGGRILPSSYSYTFMKSKTQMEV---LTNLVKKK 383
Db 521 ELIILKNDIENDSDTSEHIKRD-----RSSCKQKCEKRRRIKDEYNLRRTKIASK 575
Qy 384 NSGVYKKNFLNLPKKNKNDLDDFFKNEKEYDDLCDCRYTATIKSFLNGPAKNDVDIA 443
Db 576 PSS-DNNNSDNNNDNNNDNNND--NNDNDND--NNDNDND--NNDNDND-- 613
Qy 444 SQINVDNLRGPGCNKYKSNKE-----SWNCTGTFTNPKPGTCPPRRTQLCLR 492
Db 614 ---DNDNDNDNDNNNDNNNDNDNDLITCNMDEKHLTKIPPIIKATLDYQHAGL 670
Qy 493 TYLLHRGHEEDYKEHL-----LGASIEAQLLYKKEKDNALCALSIONSYADLA 543
Db 671 HWLLYL-----YKNNINGILADEMGLKTLQCSLLSY-----LAYEFNMGPHL 715
Qy 544 DIKGSDIK-----DYVGKMEENLNKV---NKD-----:|||:| 570
```

```
Db 716 VIVPTSILINWELKRCPCFKILSYGNQNERYKRCVGFNCKDSFHICISSYSTVVKD 775
Qy 571 ---KKR-----NEESLKIFREKWD-----ENKENVVKVMSA 599
Db 776 HLWFKRKRWKYIILDEAHNIKNFNTKRWNIILSKRDNCLLITGTPLQNSLEELSLHF 835
Qy 600 VLKXKETCKDYKQFKIPOFLRFKWDGDDFCBKRKEKYSFESFVKCECKKOD----- 654
Db 836 LMPNIFT-SHLD-----FKWFSDDL-----LAIEKSKIHHSKELIDRLHTV 877
Qy 655 -----ENTCKNKCSEYKKWIDLKKSEYKQV-DKYTKDN----- 688
Db 878 IRPIILRLKCKWKEKMPNKYEHIIKKLTRROQIILYDEFINNKNQNTLNTGNYIGLMN 937
Qy 689 ---KMYDNIDDEVKN-----:|||:|-----EANYLKEKSKCKDVNFD 720
Db 938 ILQLRKCVCNHCDFLTNKYIQTPYYWLSIRYVPVPRFFILFEKYYA-----D 985
Qy 721 DKIFNESPNEY-----EDMKCKDEIKYLNKPKTKH---DIYDIDTSDTFGDTPI 772
Db 986 FYLILFLHNEFTSLGRDVTKETSPSKGDFLAHLITGHTNLTNLYDNHHSIELYD----- 1040
Qy 773 SINANINEQSGKDT-----NTGNSSETSDSPVSHPESDAAIINVEKLSGDESSETRGIL 828
Db 1041 --NNHISELYDNHHSIELYDNHHSIELYDNHSHK-----NYKHSNGYT 1083
Qy 829 DINDP---SVTNVNVNEVHDASNTQGSVNSTSDITNGH--SESSLNRTTNAODIKIGRSGNE 884
Db 1084 YPNDPINNMMNPSGFTKTSEQFGQIVSHERDNNYHMDHNNNNLLSKEMVNSLRDDN 1143
Qy 885 QSDNQENSHSDNSGSLTIGQVPSDNTQNTVDSQNPHRDTPNALASLPDDK----- 938
Db 1144 SNNFYKSLTSNNDS-----QTSIHDNKQCDYNKL--CADTFNNINSIGNEEKRLNVL 1197
Qy 939 -----INEIGFSDSRDSENGRGDTTSTNTHDVRTNIYSERRVNSHDFIRN 984
Db 1198 NEQNNNNSKDNNNNIDNNNNIDNNNNIDNNNNIDNNNNIDNNNNIDNNH----- 1253
Qy 985 GWANNAH-----HOYITQIENN-----GIIRGOESAGNSVNYKONPKRSNFESSEN 1031
Db 1254 ---NNQHCVNNDNWPDSYPTNIHNRNAFSLKLLAQ-----NPLN-NDNNNNNNNGN 1307
Qy 1032 DHKKNQIENSRDTRKVRREIILKSKQNKCNNEYSMEYCTYSDERNSSPOPCREERKKL 1091
Db 1308 NNINMRYNRSRNSRSLNPSSTSKNS-FQLDP-LYTNSFINQDALCKN----- 1359
Qy 1092 CCQISDYCLKYNFYSIE-----YNYCIKSEIKSPEVKCFKSGQSSIPYFAGGILVVI 1146
Db 1360 -----SFFVNINIEDVHSYIYNSIYKPIPKILSFSDFELTEL-----NNNYDIL 1405
Qy 1147 VLLLSASRMGKSNEEY-----DIGESNI---EATFEENNYL----- 1180
Db 1406 SLVIDPYNRY-KSYNEYLYKMKKEGTLTNQOGLGDDINNKHIIYHKSTSENMTHMKRKTFI 1464
Qy 1181 ---NKLRSIFNQEVQETNI-----SDYSEYNYNEKNMY 1210
Db 1465 YKYNMFKVINNDTQYQNIFTDDTNNSYNSLEHNLW 1501

RESULT 37
US-10-732-923-3352
; Sequence 3352, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3352
```

Qy	879	GRSGNQSQNSQNSHSSD--NGSLTIGQVPSSE--NTQNTYDSQ-----NPHRDTPN	928
Db	1078	QNSNNK-----IEELTNVIKDLNDSIMCYKKKISBDEVKREYNEBIEIQKLNKRLNEMKDKSD	1134
Qy	929	ALASLPSDD-----KINEIEGPDSSRDS--ENGRGDTTNTDHDVRRNTIVSERRVNSHDF	981
Db	1135	KEIIKKEDEIKLKLKISNYKFEIKENYVYKSKMLNENKEBIIVDNICKE--KISSEDI	1193
Qy	982	IRNG-----MANNAHAHQVITQIEN---NGIIRGOEESAGNSVNYK	1019
Db	1194	EGKGKLNKMLLSLKCKKERSVPSVDKNKNENNELVDIVRNTYINKIEVYKKEIKONKNME	1253
Qy	1020	D--NPKSNFSSSENDHKNTQYNSRDTKRVREBIIKLSKONKCNNEYSMEYCTYSBERNS	1078
Db	1254	DLRKIIDLSNELANLKNVKNVLTEENDNLRKEM--EMKONKYNE-----NDE---	1299
Qy	1079	SPGPCSREERKKLCCQISDYCLKYFNFYSIEYNYNCIKSEIKSPYCKFKSEGOSSIPYFA	1138
Db	1300	-----ILNLDNDIIK-----LTKELSEWKERNIKLEND-----	1327
Qy	1139	AGGILVVIVLLSSARMGKSNEEYDIGESNIEATFEENNYLNKLSRIFNQ--EVOFTNIS	1197
Db	1328	-----IBQINKEYKIKENLMIKINES--INEVNVNLKNQIEIEKKCKE	1368
Qy	1198	DYSEYVNEK	1207
Db	1369	NLNTSVNWK	1378
RESULT 38			
US-10-304-095-6			
; Sequence 6, Application US/10304095			
; Publication No. US20030134275A1			
; GENERAL INFORMATION:			
; APPLICANT: Long, David M.			
; APPLICANT: Metz, Anneke M.			
; TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes			
; FILE REFERENCES: 47714-5009-US			
; CURRENT APPLICATION NUMBER: US/10/304,095			
; CURRENT FILING DATE: 2002-11-26			
; PRIOR APPLICATION NUMBER: US/09/417,485			
; PRIOR FILING DATE: 2002-06-14			
; NUMBER OF SEQ ID NOS: 49			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 6			
; LENGTH: 2184			
; TYPE: PRT			
; ORGANISM: Plasmodium falciparum			
; FEATURE:			
; NAME/KEY: unsure			
; LOCATION: (330)..(335)			
; OTHER INFORMATION: Xaa at position 330 = Leu or Ile;			
; OTHER INFORMATION: Xaa at position 335 = Asp or Gly.			
US-10-304-095-6			
Query Match 4.8%; Score 308; DB 4; Length 2184;			
Best Local Similarity 18.2%; Pred. No. 2.7e-08;			
Matches 260; Conservative 220; Mismatches 432; Indels 518; Gaps 65;			
Qy	5	FNIIPLIPLIFLVNIRINESIGRTL--YNRQDESSDISRVNSPELNNHKNHTNIDSDYE	63
Db	716	FN--YFIMPLIRPFPLTKSEQTLHKTIFFDRKIW-----NHFTKI-----	754
Qy	64	DVNNKLNISFVENKSVKGRSLSPFNKTKSYDIIIPPSYRYNDKFNLSLENEDNSGNTN	123
Db	755	--SNFCLYQIFRNKLLKGRNE-----PKMDYVQNMFNFKKGSKIKTN--	796
Qy	124	SNNFANTSEISIGNKQNTFFIQKRTLHACGKIKSI--KWICRENSSEKITVCVDPDKIQ	182
Db	797	-----KVIPIKK-----MKKXSTNKCINNFKSK---CIPKKK---	826
Qy	183	LCVANFLNGSLRTMEKFKBIFLISVNTAEKLLYNKQNEGDKPSIFCNELRNSFSDFRSSFI	242



Qy 528 ENALCSIIQSYADLADIIKSGDIIDYKGMKEENLNKVNKDKRNEESLKIPIREKXWD 587  
 Db 75 XKEVCKIINTKFTADIRDIIGTGYWDLISNRXXXXXXXXXXXXXXXXXKNDKLFDEWIK 134  
 Qy 588 ENKENVWKMVAVLNKECTCKDYKPKIPOPFLWFKWGGDDFCRKRKEKIKYISFESKVE 647  
 Db 135 VIKKDVWNVIS-WFXXXXXC-XXXXXXKIPQFFRWFSEWGGDYQDQTKMI---ETLKVE 189  
 Qy 648 CKKDCDENTCKNKSEYKWKIDLKK 673  
 Db 190 CXXXXCDDNCKSKNSYKWKISKK 215  
 RESULT 40  
 US-10-732-923-8314  
 ; Sequence 8314, Application US/10732923  
 ; Publication No. US20050108791A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edgerton, Michael D  
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
 ; FILE REFERENCE: 38-15(52796)C  
 ; CURRENT APPLICATION NUMBER: US/10732,923  
 ; CURRENT FILING DATE: 2003-12-10  
 ; PRIOR APPLICATION NUMBER: 10/310,154  
 ; PRIOR FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 24149  
 ; SEQ ID NO 8314  
 ; LENGTH: 1647  
 ; TYPE: PRT  
 ; ORGANISM: Plasmodium falciparum 307  
 US-10-732-923-8314  
 Query Match 4.6%; Score 295; DB 5; Length 1647;  
 Best Local Similarity 18.9%; Pred. No. 1e-07;  
 Matches 259; Conservative 184; Mismatches 471; Indels 460; Gaps 60;  
 Qy 16 LYNVIRI--NESIIGRTLYNRQDSSDISRVNSPE--LNNHKTNIYDSYEDVNNKLN 71  
 Db 108 LYSCLYIEENVLKGLHYEKNHNCSEHYKVEIYVNNLQWDLFFGGIENEE-- 164  
 Qy 72 SFVENKSVKKSLSFNNKTKSYDIIIPPSVYNDNFNSLSEDSNGTNSNNFANTS 131  
 Db 165 ---ELSKEMAKPYEIKNDLKCNSEFMLINEENKINSIDYLNVEKKIITYN-- 215  
 Qy 132 EISIGKDKOYTFIQKRTHLFACGKIKSIKWKICRENSSEKITVCPDRKIQLCVANFLS 191  
 Db 216 -LIIPKELK-YDLLQ-----YQEGIYMMINQEMSNVKGII----- 249  
 Qy 192 RLETMEKFEIFLISVNTAEKLLYNKQEGOPSIKVELRNSPSDFRSSFIDGDDMDFGG- 250  
 Db 250 LADEMGHKTQAITL-----ILCQKL-NKLKEIKKDESDDDHDKVGY 291  
 Qy 251 -NTD-----RVKGYINTKPSDYK-----EKVKEKLNKIKKEWKEKNAKL 290  
 Db 292 KNDDEDEKKKKKKSHMLHGDYKEDILSPEEDRTKHEKSIKEI-DIKKSY--KNSSSI 348  
 Qy 291 W---NEMVNHKGNISKECALIPAE-----EPQINLIKWNENELME 330  
 Db 349 LYCKNSTDDLUKVESDCVILIESDDTONEKVNKNKNDVNAENKID-----TENVLN 402  
 Qy 331 KRLFLNLIKCKVENKKEACFGGRLPCGSYTSFMKSKTQMEVLNLYNKKNSGVDKN 390  
 Db 403 KK--IKNIKKECVNKK-----IKNNNNNNKSNFKKN 433  
 Qy 391 NFLNDL-----FKKNKNDLDDFFNKEKYDDDLCDCRYTATIIKSFNGP 435  
 Db 434 DFINKLKGQTLIIAPAAVQWQKSEIEKFDENILNVVYHG-----NSKIISD--EEL 485  
 Qy 436 AKNDVDIASQINVDNLACFGGNYKSNNEKSNWCTGTFTNKPGTCPEPRQTLCLGRYL 495  
 Db 486 IKYDIVITSYAVB-----VNFRR-----IVNKHQPCB-----YC-GRLYL 521  
 Qy 496 LHRGHEEDYKEHLIG-ASIVEAQLKYKYEKO-----ENALCSIIQ 537

Db 522 --PNNLDIHKKYFCGPTAVRTEKLKKRKKKNTALVAMKKFDETFVPTPRNVLEIMAN 579  
 Qy 538 SYADLADIIKSGDIIDYKGMKEENLNKVNKDKRNEESLKIPIREKXWDENKENVKVM 597  
 Db 580 SKLEDEKKNKNTISN---KSLHNHNTKNGKKNQ1KQ---RNAQKDSN-----I 626  
 Qy 598 SAVLNKBTCKDYKPKIPOPFLWFKWGGDDFCRKRKEKIKYISFESKVEKCKDCDENT 657  
 Db 627 DVIVLSSDSCKE-----KSSSENS 646  
 Qy 658 CKKCKSE--YKKWIDLKSEYKQVDKYT-----KDKNKQWYDNIDEVKKEANVY 706  
 Db 647 VISPISRKTSRIIDLSNLGDKETMEQVIESLSEGKKKKNKVNKVNKSNVKS--- 703  
 Qy 707 LKESKECKDVNFDKIFNESP-----EYEDMCKCDEIKYLNKIKPKT---KHDIYD 758  
 Db 704 ---VKNRDVKWNVVKNLENFLYNVDNDCKHVLIKIYVNTNTTQIONTELEKLMGE 759  
 Qy 759 IDTFSDTFGD---GTPIS-INANINEOQSGKDSNTGNSSETSDSPVS--HEPESDAAINV 812  
 Db 760 LKVVLTNGKHIFGTQVLLNRLIVSAKYIRKELYDENTETEERKDTSGVNEEDGDTSMGV 819  
 Qy 813 EKLSGDESSSETRGILDINDPVSNTNNVNEVHDASNTQGSVSNSTSDITNGHSESLNR--- 869  
 Db 820 TNQKGDCKFKERKIKKEDEKSEIKKRKTVDVISVGGDKR-----DMSLNKRKN 870  
 Qy 870 TTNAQDIKIGRSGNEQSDNOENSHSDNSGSLTIGQVPSDENTQNTQYDQNPDRTPNA 929  
 Db 871 TTNR-----GTTMKIKGETREYASNGKNN-----NDDNNNSYYS----- 906  
 Qy 930 LASLPDDKINEIEGFDSDSRDSENGRDTTNTSHDVRRTIVSERRVNSHDFTRNGMANN 989  
 Db 907 -----DAKSYTSESSDESQYQDICKKKEPK-NTKSIRRSQKSYCKSSLESE 954  
 Qy 990 NAHQYITQIENNGIIRGQESAGSNVNYK-----NPKRNFNSENDHKK 1035  
 Db 955 -----KKEESDNSYNTDNDCCDDCCDDYNNCCDNNYCDCCDSS 996  
 Qy 1036 NIOEYNSRDTKRVREELIKLSKQKNKCN-----NEYSMEYCTYSDEBNS 1079  
 Db 997 N--SPDSYDSIIYKRS--KMSKEKKRENKMI FDESALHQIYWNRIILDRAHRIKRNRTS 1052  
 Qy 1080 PGPCREERKCLCC-----OISDY--CLKYFNFYSIEYVYCNKSEIK----- 1119  
 Db 1053 ---TTQSILMLKCGYRWCLTGTPLQNRISLIRFIEFYFYAYFCSKCKCKCLLN 1109  
 Qy 1120 -----SPEYKC-----FKSQSSIPYPAAGGILVIVILL 1150  
 Db 1110 YEMRDNKYCYFCNHSRINHFNPNKRIKPIQSGYRGEGLSGMSYL-KNEVLDKILLR 1168  
 Qy 1151 SSASRMK-----SNREYDIEGSENIATFEE--NNYLNKLSRIFN 1188  
 Db 1169 TGERKSDINLKLPIIKIRKDKLSKEKDFEYSLYKQTSQFTNTYVNSNTVLHN 1222

Search completed: November 21, 2005, 20:52:38  
 Job time : 208 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:38:54 ; Search time 8 Seconds  
(without alignments)  
170.980 Million cell updates/sec

Title: US-10-677-980-2

Perfect score: 6481

Sequence: 1 MKGYFNIYFLIPLIFLYNVI.....VQETNISDYSEYNYNEKNMY 1210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 8323 seqs, 1129788 residues

Total number of hits satisfying chosen parameters: 8323

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160.5	2.5	1315	7	US-11-077-550-141
2	160	2.5	2053	7	US-11-013-759-9
3	156.5	2.4	1169	7	US-11-077-550-20
4	151.5	2.3	1189	7	US-11-074-176-134
5	147	2.3	1420	7	US-11-077-550-110
6	147	2.3	2314	7	US-11-013-759-11
7	146	2.3	888	7	US-11-077-550-112
8	145	2.2	1992	7	US-11-013-759-3
9	145	2.2	1922	7	US-11-013-759-13
10	145	2.2	2047	7	US-11-013-759-4
11	145	2.2	2047	7	US-11-013-759-7
12	141	2.2	865	7	US-11-077-550-100
13	141	2.2	867	7	US-11-077-550-80
14	140.5	2.2	858	7	US-11-077-550-22
15	138.5	2.1	864	7	US-11-077-550-102
16	138.5	2.1	876	7	US-11-077-550-82
17	136.5	2.1	876	7	US-11-077-550-106
18	136.5	2.1	876	7	US-11-077-550-108
19	135.5	2.1	866	7	US-11-077-550-104
20	133.5	2.1	860	7	US-11-077-550-175
21	133.5	2.1	862	7	US-11-077-550-94
22	133.5	2.1	866	7	US-11-077-550-88
23	133.5	2.1	867	7	US-11-077-550-96
24	133.5	2.1	867	7	US-11-077-550-98
25	133.5	2.1	870	7	US-11-077-550-92

26	133.5	2.1	871	7	US-11-077-550-84
27	133.5	2.1	871	7	US-11-077-550-86
28	133.5	2.1	871	7	US-11-077-550-90
29	132.5	2.0	862	7	US-11-077-550-171
30	130.5	2.0	873	7	US-11-077-550-167
31	130	2.0	438	7	US-11-074-176-152
32	129.5	2.0	871	7	US-11-077-550-2
33	129.5	2.0	871	7	US-11-077-550-8
34	129.5	2.0	871	7	US-11-077-550-26
35	129.5	2.0	871	7	US-11-077-550-153
36	129.5	2.0	873	7	US-11-077-550-6
37	129.5	2.0	873	7	US-11-077-550-149
38	129.5	2.0	873	7	US-11-077-550-151
39	129.5	2.0	879	7	US-11-077-550-159
40	129.5	2.0	894	7	US-11-077-550-4
41	129.5	2.0	1013	7	US-11-077-550-18
42	129.5	2.0	1127	7	US-11-077-550-40
43	129.5	2.0	1129	7	US-11-077-550-42
44	129	2.0	866	7	US-11-077-550-32
45	129	2.0	876	7	US-11-077-550-66

#### ALIGNMENTS

##### RESULT 1

US-11-077-550-141  
; Sequence 141, Application US/11077550  
; Publication No. US20050244435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 141  
; LENGTH: 1315  
; TYPE: PRT  
; ORGANISM: Clostridium tetani  
US-11-077-550-141

Query March 2.5%; Score 160.5; DB 7; Length 1315;  
Best Local Similarity 18.0%; Pred. No. 0.0041;  
Matches 264; Conservative 178; Mismatches 457; Indels 565; Gaps 69;  
Qy 32 YNRDESSDLSRVNSPELNHNHKTNIYDSVDVNNKLNLSFVENKSVKKKRSLSFINNK 91  
Db 77 YLRDTSKDFLQTNVKKLFNRKNNVAG---EALDKLIIN-----ALPYLGN- 120  
Qy 92 TKSYDIIPPSYSYRNDKF-----NSLSEN---EDNSGNTNSN-----NFPAN 129  
Db 121 --SYSLI-----DKFDTSNSVSVFNLLQDPDPSGATTKSAMLNTNLIIFGPGVNLNKE 170

```
QY 130 TSEISIGKDNKQYTFIOKTHLPACGIRKRSIKWICRENSKIVCVDPDKIQLC-----184
Db 171 VRGIVLVRDNKY-----FPCRDGFGSI-----NMQAFCPPEYVP 204
QY 185 -----VANFLNSRLTMEKFEIPLISV-----TEAKLLYNKE-----219
Db 205 TFDNVIENITSLTIGSKSYQDPALLMHELHVLHGLYGNQVSSHIIIPSKQEIYMOHT 264
QY 220 -----GKDPISFCNELN-----SFGDMDPGGNTDRVKG 257
Db 265 YPISAEELTFGGDANLISIDIKNDLYEKTLDYKAIAKLSQVTSNDPNIDIDSYKQ 324
QY 258 YINTKPSDYKERKVEKLNKIKKWEKKNKANLWMI-----VNHKGNISKECAII- 309
Db 325 IYQOKY-QFDKDSNGQYIVN-----EDKFIILYNSIMYGTETELGKFKFNKTRLSYFS 377
QY 310 ----PAEPEQ-INLWIKEMENFLMEKKRL-----FLNKKDKCVENKYE 349
Db 378 MNHDPVKIPNLDDTIYNDTEGFNIESKDLSEYKGMQMRVNTNAPRNVGSGVLVKLIG 437
QY 350 ACFGCGRLPCSSYTFMKSKTQMEVLNTNLYKKNSGVD-----KNFLNDLFKK 400
Db 438 LC-----KKLIPTNIRENLNRTASLTDLGGELCIKIKNEDLTFIAEK 482
QY 401 NKNDLDDFPKNEKEYDLDL-----CRYTATIIFKSLNGPAKNUVDIASQINV-ND-- 450
Db 483 S-----FSBEPQDEIVSYNTKKNPLNFNYSLDKIIV-----DYNLQSKITLPNDRT 529
QY 451 ---LRG--FGCNYSKNEKSNCTGTFTNKPFGTCEPPRQTLCLGRTYLLHRGHEBDYK 505
Db 530 TPVTXGIPYAPEYKSN-----AASTIEIH-----553
QY 506 EHLGASIIYEAQLKYKKEKENALGSIIONSYADLADIIGSDIIDKYGKKMEENLN 565
Db 554 ----NIDDNWIIYQVLYAOKSPITLQRIITWNSVD--DALINSTKIYSYF----PSVIS 601
QY 566 KVK-----DKRNEESLKIFREKWD-----587
Db 602 KVNQGAQGLIFLOWVRDIIDFTNESSQKTTIDKISDVSTIVPYIGPALNIVKQYBGNF 661
QY 588 -----ENKENVMKMSAVLKNKETCKDYDKFQ 614
Db 662 IGALETTGVVLLBYIPEITLVPVLAALSABSSQKEKIITIDNFLE-----KRYEKWI 716
QY 615 KIPQFL--RMFKWGGDDFCERKKEKIYSFESFKYCKKXDCDENTCKKCYKWIIDLK 672
Db 717 EYVKLVAKWLGTVNTQF-QKRSYQMYRSLBYQVDAIKKIID-----YEYKIYSGPD 767
QY 673 KSEYKQVDKYTKDNKKKQYNDIDEVK-----EANYLKEKSK-----ECKD--VN 718
Db 768 K--EQIADEINNLK-----LEEKANKAMININIFMRESSRFLVNQMINAEKQOLLE 819
QY 719 FDDKIFNESNEYEDMCKKCDIEIKLANEIKYPTKHDIYDIDTSDTFGDTPLISINANI 778
Db 820 FD-----TQSKNLMQYIKANSKPTGITELKKLESK-----INKVFTPIPPSYSKNL 867
QY 779 NEQSGKDTNNTGNSSETSDSPVSHPEPSDAINVEKLSGDESSSETRGILDINDPSTVNN 838
Db 868 -----DCWVDNEEDIDVLKKSITLN-----LDIN-----ND 894
QY 839 VNEVDASNTQGSVNTSDI-----TNG-----HSESSLNRTNNAQDI-----876
Db 895 I--ISDISGFGNSSVITYPDQALVPGINGKAIHLVNNESSEVIVHKAMDIEYDNMNFNTV 952
QY 877 -----KIGRSGNEQSDNOSNSSHSDNSGLTIGQVPSBDNTQNTY-----DS-----919
Db 953 SFWLVRVKVASHLEQVGTNEYISIISSMKHSLSIGSGWSVLSKGNLIIWTLKDSAGEVR 1012
QY 920 QNPHRDTFNALASPDDKINEIEGFPDSDRSENGRGTDTTNTHTVRRRTIVTSERRVNSH 979
Db 1013 QITFRDLP-----DKFNAY-----LANKW 1031
```

```
QY 980 DPTIRGMANNNAHQYITQIENNGIIRGQESAGSNVYKDNPKRSNFSSENHKKNIQE 1039
Db 1032 VFITITNDRLSSANLYI-----NGVLMSGABITGLGAIREDN-----NITKLDRCNMNNQ 1082
QY 1040 YNSRDTKRVREETIIKLSKQNKCNNEY-SMEY-----CTYSDERNSSP-GPCSRERK 1089
Db 1083 YVSIDKRFPCAKALNKEIEKLYTSVLSITFLDFWGNPLRYDTEYVLIIPVASSKDVQL 1142
QY 1090 KLCQCSIDVCL-----KY-----FNFYSIEYVNCIKSEIKSPEYKCFKSEGOSSIPYPAAG 1140
Db 1143 K---NITDYMVLNAPSNTNGKLNIVYRRLYNGLKPIIK-----RYTPNNEIDSFVKS 1192
QY 1141 GILWVILLSSASRMKSNEEYDIGESNIEATPEENNYLNKLSRIEN-----Q 1189
Db 1193 GDFIKLY-----VSYNNEHIVGYP-----KGNAPFNLDRIILRVGNAPGIFLYK 1238
QY 1190 EVQETNIDSYSEYN-----YNEKN 1208
Db 1239 KMEAVKRLDLKTVSVQLKLYDDKN 1262

RESULT 2
US-11-013-759-9
; Sequence 9, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE OF INVENTION: PROTEIN OF MORAXELLA
; FILE REFERENCE: 1038-92LMIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; PRIOR FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361,619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-9
```

```
Query Match 2.5%; Score 160; DB 7; Length 2053;
Best Local Similarity 18.1%; Pred. No. 0.0074;
Matches 218; Conservative 149; Mismatches 446; Indels 390; Gaps 54;

QY 33 NRQDESDISRVSPELNNHKTNIYDSYEDVNNK--LINSFVENKSVKKRSLFINN 90
Db 643 NNGDLVDSGNIATPTTYNISVKTTKLNSNGTSGNNKFSVSNADNNSLVTAKDLADLYNK 702
QY 91 KTKSYDIIPSYVRN--DKFNSISENDSNGT-----122
Db 703 VNETAOSALPSFKVQNGDNNNAITVGKDTNGKTFNTLKLKGENVNITNRTATGTTFG 762
QY 123 -NSNNFANTSEISIGKDNKQYTFIOKTHLPACGIRKRSIKWICRENSKIVCVDPDKI 181
Db 763 IDQSNGLTTPKLVGSDTNGNRLVIEQVP----SADGNSTKNIKGLSPTLPSPSPGR 818
QY 182 QLCVANPLNRLTMEKFEIPLISVNTAKLIYN-KNEGKOPSIIFCNELRNSPSPDRSS 240
Db 819 NIALGN-----TIEKDKSNAASIDDLVNLNAGFNKNGKD-----KDFVST 859
QY 241 FIGDDMDP-CGN-----TDRVKGYNTEKFDYKCKVKNVKNLNIKEWKEKN 286
Db 860 Y--DTVDFIDGNATTATVTVDENQKTSKAYDVNV-----DEKTIELTGDKGKQLGVK 911
QY 287 KANLWNHMIYVNHKGNIS-----KECAIIPAEFPQINLWIKEMENFLMEKKRFLNKK 341
Db 912 TIKLTE--TSTNGNATFTSTDDHALVKASDIAGNL-----NTLAEE-----951
```

```
Qy 342 CVENKKEACFGGCRLLPCSSYTSFMKSKSTQMEVLNLYKKKSGVDKNNFLN----- 394
Db 952 -IHTTKGTA-----NTALQTFVKKVDENDKADDTNALTGVKOGTS 991
Qy 395 ---DLFKKKNNDLDDFFKNEKEYDDLCDCRYTATIIKSFNGPKANDVDIASQINVD- 450
Db 992 GKVTNLKLGKNGLD-----IKTDXDGVTFGINTQSGLKAGDS 1030
Qy 451 --LRGFGCNYSNNEKSNWCTGTFTNKF-----PGTCEPPRRQTLCLGRTYLL 496
Db 1031 TTLANNGLSIXNTASNEQIQVAGDVKPFAMVNGVVGAGIDGTRITRDEIGFTGT----- 1086
Qy 497 HRGHEDYKHLGASIVYEAQLLYKKBKDNALCSTIIQNSYADLADIHKSGDIIKDY 556
Db 1087 -NGSLDKSKPHLSKDG- -NAGGKKITNIQSCE-----IAKNSH-----DAVTGGKI-----Y 1131
Qy 557 GKMEENLKNYKDKKRNEESLKIIFREKXWMDENKENVKWSAVLNKTKETCDYDKFKQI 616
Db 1132 DLKTELE-NKISSTAKTAQNSLHEFSVA--DEQGNF-----TVSNPYSSYDTSKTSDV 1182
Qy 617 PQFLRFKMGDDDFCEKREKIKYSFSPKVECKKDCDENTCKNKCSE--YKKWIDLKKS 674
Db 1183 ITP-----AGENGITTKVNGVVRVIGIDQTKG 1209
Qy 675 EYEKQVDKYTKDNK-----KMYDNDIDEVKNKEANVYLKESKECKDWN 718
Db 1210 ---LTPPKLTVGNNGKIVINSQNGWTITGLSNTLANVTNDKSGVRTEQGNIIKD-- 1264
Qy 719 FDDKI-----FNESPN-EYEDMCKKDBEIKYLN-----BIKY---PKTKHDI 756
Db 1265 -BDKTRAASIVDVLSAGFNQNGEAVDFVSTYDTVPFANGNTTTAKVYDTSKTSKV 1323
Qy 757 YDIDTFSDTFDGTPIST-----NANINEQOSGKDTSTNGSETSDSPVSHEPESDA 808
Db 1324 YDNNV-----DDTIEVKKLGKVKYTTLTSTGTGANKFALSNQATGDLVK---ASDI 1374
Qy 809 AINVEKLSGD-----ESSSETGILINDPSV--TNVNEVHDASNTQSGVNTSDI- 858
Db 1375 VAHLNLTLSGDIQTAAGSQAANSAGYDADGNKVIYDSTDNKYQAKN-DGTVDKTKEVA 1433
Qy 859 -----TNGHSESSLNRTNAQDIKIGRSNEQSDNQENS-----SHSSN---SG 900
Db 1434 KOKLVAQAOTPDGTQAQNVKSVINKSQVNDANKKQGINEDNAFVKGLEKAASDNKTNA 1493
Qy 901 SLTIGQVPESENTQNTYDSQNPHTDTPNALASLP-----SDDKINEIEGF 945
Db 1494 AVTVGDNLNAVQPLTFAG-----DTGTTAKKGGETLTIKGGQDTPNKLTDNNIGVAGT 1548
Qy 946 DSSRDSNGRGDTTSTNTHDVRRTNVS-----ERRVNSHDFIRNGWMANN----- 990
Db 1549 D-----GFTVKLAKDLTNLSVNAVAGTKIDBKGISFVD--ANGQAKANTPVLNANG 1597
Qy 991 -----AHQYITQIEN-NGIIRGOEESAGNSVNYKDNPKRSNPFSS 1029
Db 1598 LDGKGKVISNVGKGTQDTPAANVQQLNEVRNLLGL--GNDNADGNQVNIADIKKDPNSGS 1655
Qy 1030 END 1032
Db 1656 SSN 1658
```

## RESULT 3

```
US-11-077-550-20
; Sequence 20, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
```

```
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR FILING DATE: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 20
; LENGTH: 1169
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; US-11-077-550-20
```

```
Query Match 2.4; Score 156.5; DB 7; Length 1169;
Best Local Similarity 18.0; Pred No. 0.0062;
Matches 220; Conservative 169; Mismatches 415; Indels 421; Gaps 55;
Qy 252 TDRVKGINTKFSYKQKVKELNNIKK-----EMWEKKNKANLWNHIMVNHGNIKECA 307
Db 40 TDRI-WIIPERYTGYKPEDFNKSSGIFNRDVCYVDPYLN-----TNDKKNI----- 87
Qy 308 IIPAEERQINLWIKENWENFLMEKKRLFLNATKDCVENKKEACFGGC-----RLPSS 361
Db 88 -----FLQTMIKLFNRKSKPLGKLEMIINGIPYLGDRRVPLE 128
Qy 362 YTSFMKSKTKQMEVLNLYKKKSGVDKNNFLNDFKKNKNDLDDFPKNEKEYDDLCD 421
Db 129 FN-----TNI-----ASVTNKLISNPGEVERKKGI-----FAN----- 157
Qy 422 RYTAIIKSFNGPKAND---VDIASQINVNDLRFSG-----C-NYKS--NNEKSWNC 468
Db 158 -----LIIFGPGVLENENETIDIGIQNHPASREGFGIMQMKFCPEVSVFNNVQE--- 208
Qy 469 TGFTNKFPGTCEPPRQTLCLGRTYLLHRGHEDYKHLGASIVYEAQLLYKKEKDE 528
Db 209 ---NK-----GASIFNRGRGYFSDPALILMHLELHVHLGL-YGIKVD- 246
Qy 529 NALCSIIIONS---YADLADIHKSGDII-----KDYVGKME-----ENL 564
Db 247 ---LPIVFNKKPFMQSTDAIQAEELTYFGQDPSIITPSTDKSIYDKVLQNFGRIVDRL 303
Qy 565 NKVNKDKKRNEESLKIIFREKXWMD-----ENKENVKNVMSAVLNKNETCKOYDKFKQIPQF 619
Db 304 NKVLVCISDPNININIVKNKFKDKYKFVEDSEGKYSI-----DVESFDKLYKS 351
Qy 620 LRW-FKEW--GDDPCEKREKIKYSFSPKVECKKDCDENTCKNKCSEYKKNWIDLKSEY 676
Db 352 LMFGFTETNIAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYTIIEGPNISDKDM 405
Qy 677 EKQVDKYTKDNKKNWMDNIDEVKNKEANVYLKESKE-----CKDYNFDDKINESPNE 730
Db 406 EKEYRGONKAINQAV---BEISKEHLAVYKIQCKSVKAPGICIDVDNEDLFTIADKNS 462
Qy 731 YEDMCKKDBIKYKPKTKHDIYDIDTFSDTFGDTPIISINANINEQOSGKDTSTNT 790
Db 463 FSDLSKNERIEYNTQSNYIENDPPINELIILDTL-----ISKLELSENTES 510
Qy 791 GNSETS DSPVSHEPESDAAINVEKLSGDESS-----SETRGILINDPSVNTNNEVHD 844
Db 511 LTFDNVDVPVVEKQPA-----IKKIFTDENTIFOYLYSQTPP-LDIRDISLTSFDDALL 564
```

```
Qy 845 ASNTQGSVNTSDITNGHSESSLNRTTNAODIKIGRSG-----NEQSDN-----QENSSHS 895
Db 565 FSNKVSFFSMDYI-----KTAN-----KVYEAGLFAGWVKQIVNDFVIEANKSNT 610
Qy 896 SDNSG-----SLTIGOVPSEDTQNTYDSQN-----PHRDTPNALASIPS- 935
Db 611 MDKIADISLIVPYIGLALVGNETAKGNFENAFIAGASILLBEIPBLILPVVGAFLLES 670
Qy 936 --DDKINEIFGSSRDSRSGD-----TTSNT--HDVR----- 966
Db 671 YIDNKNKIITIDNALTKREKSDMYGLIVAOWLSTVNTQFVTKBGMVKALNYQAAL 730
Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHHOYITOIEN--NG----- 1003
Db 731 EEIKYRNIYSEKESKINIDFNDSKLNNEGQINQAINNPNFINGCSVSYLMKKMPL 790
Qy 1004 ----ITRGEESAGNSVNYKDNPKRSNFSSENDHKNIQOYNSR-----DTRKVRBEIIL 1055
Db 791 AVEKLLDPDNTLKGULLNYIDENKLYLIGSAEYKSKVNYLKTIMPFDLSIYNTDILI 850
Qy 1056 SKQKNCNE-----YSMEYCTYS-----DERNSSPGPCSRERKK 1090
Db 851 EMENKYNSEILNLIILNRKYDNKLLDLSGVGAKEVYDGVGLNDKNQFKLTSSANSKIR 910
Qy 1091 LCQISDYCLKYRNFYIEYNYCNCKSEIKSPEYK-----CFKSEG----- 1130
Db 911 VTQONIIFFSVLDFSVFW-----IRIPKYNDGIONYIHNEYTIINCMKNNSGWKI 964
Qy 1131 -----QSSIPYPAAGGILWVIVLLLSA-----SRMGK 1158
Db 965 SIRGNRIIWLIDINGTKSVFFEYNIREDISEYINRWFFVTYNNLNNAKIYNGKLES 1024
Qy 1159 SNEYDICE--SNIEATFENNLYN-----KLSRIFNOEVOETNISD-----YSEY- 1202
Db 1025 NTDIKDIREVTANGELIIFKLDGDIRTOFTWMKYFSIFNTELSQNSIEERYKIQSYEYL 1084
Qy 1203 -----NYNEKNMY 1210
Db 1085 KDFWGNPLMYNKYMFNAGKNSY 1109

RESULT 4
US-11-074-176-134
; Sequence 134, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-134

Query Match 2.3%; Score 151.5; DB 7; Length 1189;
Best Local Similarity 19.2%; Pred. No. 0.012;
Matches 159; Conservative 138; Mismatches 290; Indels 243; Gaps 41;

Qy 513 IYE--AOLLKYK-YKEKDENALCSIIQNSVADLADIIKSGDIINDKYGKMEENLANKVVK 569
Db 160 IFEEAAGVLHFQKQKEEAQNL-----KKTQDNLIRIND 193
```

```
Qy 570 DKCRNEESLKIIFREKWMDEKENVKVMGAVLKNKTKCDYDKFQKIPQFLRWFKWEGDD 629
Db 194 LVKELESRLPLENEQ-----SSLAKY-KFQK-----S 220
Qy 630 FCEKRKEKIYSPSPFKVECKKKDCDENTCKNK-----CSEYKKWIDLKSEYEKQV 680
Db 221 GLDKKLLSLAFIEIENIQOREDIQKSAADNKKILLAKLDDDEVKDSQAAVTQKRAEYKCLR 280
Qy 681 DK--YTKDNKKMYMDIDEVKNKEANVYLKESKECKDVNFDDKIFNESPEYEDMCKKC 738
Db 281 DERDHTQNKLLKSLSEL--NASLQMAEQSRQ-----FDDATKEEYKNQYKQLAQNL 332
Qy 739 DEIKY-LNEIKYPTK-----HDIYDITDFTDTPGD--GTPISINANI-----NEQSGKD 786
Db 333 VQLKADLDELKKEKKLQDEQDVLKIERGOLT--GELNEDPEELNKKLLDIRNNYMLQLQD 391
Qy 787 TSNITGNETSDSPVSHPEPESDAAINVEKLSGDESSSETRG-----ILDINDPSTVNNVNE 841
Db 392 QATNNQIVNLNSDLRRSQADTTYQ-----TGDVSKOLTDAQOLEQRLEGKLLTDKQK 447
Qy 842 VHDA---SNTQGSVNTSDITNGHSESSLNRTTNAQ--DIKIGRSGNEQSDNOENS--- 892
Db 448 EQNAIVRINKQNQLNT-ELTN-----LRQVVNAERNELEKVEARHEALVNIQKRHEGY 500
Qy 893 -----SHSDNSGSL-TIGQV-----PSEDNTQNT-----YDSQNPHRDTPNA 929
Db 501 YGVRNVNLHNDFAVIGAVGELITFPAELEAAMTALGGVQDGLITESRISARNAINK 560
Qy 930 LAS-----LPSD-----DKINELEGDSR-----DSENGRGDTTSNTH-- 963
Db 561 LKQNHGRATFLPDGLRQYGIPOQSTVTTLLKSYDGFGRGIASDLVESKTDQDITAAINYL 620
Qy 964 -----DVRTNTIVSRRVNSHDFI-----RNGMANNAAHOYITQIENNGI 1004
Db 621 GSVVIVDTIDTAMSAQVRNRYRIVTLDGDIVSPGSGMTGGQNRQNSPLQATATEINQL 680
Qy 1005 IRQGEESAGNSVNYKDNPKRSNFSSE-----NDHKNIQEYNSRDTKRVREIILKLSQNK 1060
Db 681 EK-QIKTLKQNLN-EDQDKLENLVDQSNKVNABEQDLQDALRETSQAINEAAISFQGEK 738
Qy 1061 -----CNNEYSMEYCTYSD-----ERNSSPGCSR--EERKKLCCOISDYCLK 1101
Db 739 EVKRLDANTLYKSRIKDRNDRIELKQIKKEANDKMLTKOGEEOAKWDLQDK-IK 797
Qy 1102 YFNFYSIETNYCNKSEIKSPYKCFKSEGOSSIFYFAAGGILWVIVLLLSASRGMKSNE 1161
Db 798 NFNNLS---QRIQDELKSLDPK-----IAVVTNKLNLSSQENKKNH 836
Qy 1162 EYDIGESNIR-----ATFENNLYNKLRIFFNOEVOETNISDYSEYNYNE 1206
Db 837 QIDNSEKQIEDLTKULTILAQDENSEMNTANLEKQKSTI-----EQKNNE 882

RESULT 5
US-11-077-550-110
; Sequence 110, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
```



```
QY 124 SNNFANTSEISIGKDNQYTFIQKTHLPACGIIKRSIK-----WICRENSE 170
Db 1174 SNTLANTNDGAG-----HALSQGLANDTDKTRAASIGDVLNAGFNLOGE 1220
QY 171 KITVCVPDRKIQLCVANFLNSRL---ETMEKFEIFLISV-NTEAKLLYNKEGKPSIF 226
Db 1221 AVDFVSUYDIDVDFIDGNATTAKVYDDTSKTSKVYDVNDNKTIETVTSKKLGKVTTL 1280
QY 227 CNELRNSFSFRSFI GDDM-----DFGNTDRVKGYINTKFSYDIYKEKNVEKL 275
Db 1281 TKTSGANGNATKFSAADGALVKASDIATHLNTLAGDIQTAKGASQASSASYVDADGNKV 1340
QY 276 --NNIKKEWEKKNKANLWNHMI VNHKGNISKECAIIPAEBPQINLWIKENWENFLMEKR 333
Db 1341 IYDSTDKKYYQ-----VNDKGQVDKNKEV-----1364
QY 334 LFLNIDKVCVENKKEACFGGCRLPSCSSYTSFMKSKTQMEVLNLYKKNSGVDKKNFL 393
Db 1365 ----AKDKLVAQ-----AQTPDGTLAGMNVKSVINKEQVNDANKQ--GINEDN-- 1407
QY 394 NDLPKXNNKNDLDP- FKNEKBYDDLCDCRYTATIIKSFNLGPAKNDVDIASQINVDLR 452
Db 1408 --AFIKGLENAAKDTKKN-----AAVTVGDLNVAQAQTLPTFAGD----- 1445
QY 453 GFGGNYKSNNEKSNWCTGTFTNKFPGCEPPROTLCIGRTYLLHRGHEDEYKEHLLGAS 512
Db 1446 -----TGTAKK-----LGETLTIKGG-----1462
QY 513 IYEAQLLYKKEKENALCSIIQNSYADLADIIGSD-----LIKDYVGKMEENLKV 567
Db 1463 -----QYDTNKLTD-----NNGVVGAGDTGTVKLAKOL-----TNLSV 1497
QY 568 NKDKRNEESLKI PREKWDENKENVWMSAVLNKNETCKDYDKPQKIPQLRWFKEWG 627
Db 1498 NAGGTRIDEXGISP-----VDANGQA--KANTPVLSAN-----G 1529
QY 628 DDFCEKEKEKIYSPESFKVECKKDCDENTCKNCSEYKWIIDLKSEYKQVDKYTKDK 687
Db 1530 LDGGKRISNIGA-----AVDDNDAVN-----FKQFNEVAKT-----1561
QY 688 NKMYNDIDEVKNKEANVYLKEKCEKQVDNFDKIFNESPEVEDCKKDEIKY--L 744
Db 1562 -----VNNLNQNSGASLPFVVTDTANGKPI NGTD-----GKPKAIKAGDKYHYANA 1610
QY 745 NEIKYPTKHDYIDIDTFSPTFGDTPISIN-----ANINEQSQGKOTS 788
Db 1611 NGVPVDKDGKPI TDADKLANLAHAGKPLDAGHQVVASLGGNSDAITLTNIKSTLPQIDTP 1670
QY 789 NTGNSETSDSPVSHESPDAAINVEKLSGDESS--SETRGILLDINDPSVTNNVEVH--D 844
Db 1671 NTGNANAGO-----AQSLPSLSAAQQSNAASVKDVLNFGNLQTNH-NQVDFVK 1718
QY 845 ASNTQGSVSNT-SDITNGHSESSL--NRKTTNAQDIKIGRSGNEQSDNQENSSHSDN--- 898
Db 1719 AYDTVNFVNGTGADITSRSADGTMSHITVNTALAAATDDGDNVLKAKDGKFKYADLMP 1778
QY 899 SGSLTIGQVPSEDNTQNTYDSQPHR---DTPNALA-----SLPSDDK 938
Db 1779 NGSLKAGKSASDAKTPGTGLSLVNPAGNAGSGTGDAVALNNLSKAVFKSKDGTITTTVSSDG 1838
QY 939 INEIEGFD-----SRSENGRGTTN-----THDVRTNIVSRRVNSHDFTRNGMANN 989
Db 1839 IS-IQKDNSSITLSKDLGNVGGKVI SVNGKGTXTDAAVY--QQLNE---VRNLGLG 1891
QY 990 NALHOYTQIENNGIIRGOESAGNSVNYKDNPKRNSFSEND 1032
Db 1892 NA-----GNDNADGNQVNIADIKDPNSGSSN 1919
```

RESULT 7

US-11-077-550-112

; Sequence 112, Application US/11077550

```
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Poster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; PRIOR FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 112
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; US-11-077-550-112
```

Query Match 2.1%; Score 146; DB 7; Length 888;

Best Local Similarity 18.7%; Pred. No. 0.019;

Matches 182; Conservative 138; Mismatches 344; Indels 308; Gaps 45;

```
QY 252 TDRVKGYINTKFSYDIYKEKNVEKLNLIK-----EMWEKKNKANLWNHMI VNHKGNISKECA 307
Db 51 TDRI-WIIPRYTFGVKPEDFNKSSGIFNDRVCEYDPPDLN-----TNDKKN-----98
QY 308 IIPAEBPQINLWIKENWENFLMEKKEKRLFLNIKDKCVENKKEACFGG-----RLPCSS 361
Db 99 -----FLQTMILFNRIKSKPLGEKLEMIINGIPYLGDRRVELEE 139
QY 362 YTSFMKSKTKQMEVLNLYKKNSGVDKNNFLNDLPKNNKNDLDDPFKNEKEYDDLDCD 421
Db 140 FN-----TNI-----ASVTVNKLISNPGEVERKGI--FAN-----168
QY 422 RYTATIIKSFNLGPAKND---VDIASQINVNDLGRFG-----C-NYKS--NNEKSWNC 468
Db 169 -----LIIPGPGVLNENETIDIGIQNHFPASREGFGIMQMKFCPEYVSFVNNVQ--219
QY 469 TGTFTNKPCTCPPTROTLCIGRTYLLHRGHEDEYKEHLLGASIEQAQLLYKKEKDE 528
Db 220 -----NK-----GASIFNRRGYFSDPALILMHELIHLHGL-YGIRKVD--257
QY 529 NALCSTIONS---YADLADIIGSDII-----KDYVGKME-----ENL 564
Db 258 ---LPIVPEKKYFMOSTDAIQABELYTFGQGDPSIITPSTDKSIYDKVQINFRGIVDRL 314
QY 565 NKYNKDKRNEESLKI PREKWD-----ENKENVWMSAVLNKNETCKDYDKPQKIPQF 619
Db 315 NKVLVCISDPNININIVKNKPKDKYKFVEDSEGKYSI-----DVESFDKLYKS 362
QY 620 LRW-FKEW--GDDFCRKRKEKIYSPFSFKVECKKDCDENTCKNCSEYKWIIDLKSEY 676
Db 363 LMFGFTETNIAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYTIIEGFINSDKM 416
QY 677 EKQVDKYTKDKKNKMYNDIDEVKNKEANVYLKEKSK-----CKVDNFDKIFNES 727
Db 417 EKEVRGQNKAINQAV---EISKEHLAVYKIQMCKLSDDDDKAPGICIDVDNEDLFIAD 473
```

Qy	215	YKNNEGKOPSI-FCNELRNSFSDRSFICDDMDFGGNTDRVKGYINTKESDYKERNVE	273
Db	676	-----NGSNPTGIANTAR-----ITRDKIGFAGS-----DGAVDT-----	706
Qy	274	KLANNIKKEWWEKKNANLWHMIVNHKGNISKBCAIIPAESPQNLWIKENENFLMEKKR	333
Db	707	-----NKPYLDDQKLGQGNVKITNTGINA-----	730
Qy	334	LFLNIKDCKVENKCYEACFGGCLPCSSYT--SFMKSKTOMEVLTNLY--KKKNSGVDKN	390
Db	731	-----GKKAITGLSPILPAIOSSRNIEHLNTIQDKDKSNAASIN	771
Qy	391	NFLNDLFFKKNKNVDLDDPFKNEKEYDDLDCRYTATIIKSFPLNGPAKNVDYIASQINVD	450
Db	772	DIILNTGFNLKNNNPIDFVSTYDIDVFANGNATTATVTHDTANKTSK-----VYDVNVDD	827
Qy	451	LRFGFCNYKSNKESWNCTGTFNKEPGTCEPPRQTLCLGR-----YLLHRGHBE-----	502
Db	828	-----TTIHLTGTDDNKKLGVKVTKTLNKTSAANGTATFNVNSSDEDALVNA	874
Qy	503	DYKEHL--LGASIIY-----EAQLLYKYEKENALCSIIIONSVADLADIILKSGDIIK	553
Db	875	KDAENLNTLAKSIHTKTGADTALQTFVVKVDE-----NNNADDANAIT-----	920
Qy	554	DYYGKMEENLKNVKDKKRNEESLKI PREK-----WMDENKENVKVMASVL-----K	602
Db	921	--VGQKNAN--NQVNTLTLKGENGLIKTDKNGTVTFGINTTSGLKAGKSTLNDGGLS	976
Qy	603	NK-----ETCKDYDIPQKI-----POPLRW	622
Db	977	NPTGSBQIQVGADGVKFAKNNVNVGAGIDGTTTRITRDEIGFTGNSLDKSKPHLSK-	1035
Qy	623	FKEWGDDPFCKRKEKLYSPESFKVECKKDC-----DENT--CKNKCSEYKKWIDLKKS	674
Db	1036	-----DGINAGGKLTINQSGEIAONSHDAVTGSKIYDLKTELENKISSTAKTAQNSLH	1089
Qy	675	EY-----EKQVDKYTKDNKKMYD-----NIDEVK-----	699
Db	1090	EFPSVADQGNFTVSNPYSSYDTSKTSDDVITPAGENGITTKVNGVVRVIGIDQTKGLTTP	1149
Qy	700	-----NKEANYLKEKSKCKDVNFPDDKI-----	723
Db	1150	KLTVGNNGKGIVIDSQNGQNTITGLSNTLANVTNDKGSVRTTEQGNIIKD-----EDKTRA	1206
Qy	724	-----FNESPN-EYEDMCKCKDEIKYLN-----EIKY---PKYKHDYIDDTFS	763
Db	1207	ASIVDVLNAGFNLQNGEAVDFVSTYDYNFADGNATTAKVYDTSKTSKVYDVNVV--	1264
Qy	764	DTFGDGTPISI-----NANINEQQSGKDTNTSGNSETSDSPVSHEPESDAAINVEKL	815
Db	1265	-----DDTTIEVKDKLGVKVTITLTSTGTGANKFALSNQATGDLVK---ASDIVAHLNTL	1317
Qy	816	SGD-----ESSSETRGILDINDPSV--TNVNVEVHDASNTQGSVNTSDI-----	858
Db	1318	SGDIQTAKGASQANNSAGYVDADGNKVIYDSTDNKYQAKN-DGTVDKTEKAVDKLVAQ	1376
Qy	859	--TNHGSBSLNRNTNAQDIKIGRSNEGSDQNS-----SHSDN-----SGSLTIQGV	907
Db	1377	AQTPDGTLAQMNVKSVINKEQVNDANKQGINEDNAFVKGLEKAASDNKTQNAAVTVGDIL	1436
Qy	908	PSNDNTQNTYDSQNPHRDTPNALSLP-----SDDKINELEGFDS-----	947
Db	1437	NAVAQTPLPFAG-----DTGTTAKLGETLITIKGGQTDNKLTDNNIGVVAGTGFVKL	1491
Qy	948	SRDSEN-----GRGDTTINTHVRRTNIVISERRVNSHDFIRNGM-----AN	988
Db	1492	AKDLTNLNSVNAAGTKIDDKGVSFVDSGQARANTPVLNANGLDLGKGVISNVGKGKDT	1551
Qy	989	NNAHHQYITQIEN-----NGIIRGQBEESAGNSVNYKONPKRNSPNSEND	1032
Db	1552	DAANVOQLNEVRNLLGLGNA--GNDNADGNQVNIADIKKDPNSGSGSSN	1597

859	QY	--TNGHSBSLLAKRTINAQDIKLGKSGNEQSDQENIS-----SHSSDN-----SGLLTIGQV	907
1377	Db	AQTPODGTLAGMNVKSVINKEQVNDANKKQGINEDNAFVKGLEKAAASDNKTKNAAVTVGDL	1436
908	QY	PSEDNTQNTVDSQNPHRDTPNALASLP-----SDDKINELEGFDS-----	947
1437	Db	NAVAQPLTFAG-----DRGTTAKKLGELTIKGGQDTNKLTDNNIGVVAGTGTGTVKL	1491
948	QY	SRDSEN-----GRGDTTSTNDHVRTNIVSERPVNSHDFIRNGM-----	988
1492	Db	AKDLITLNSVWAGGTIKDDKGVSFVDSGQAKANTPVLSSANGLDLGGKVISNVGKGTGKDT	1551
989	QY	NNAHHQYITQIEN-----NGIIRGPESAGSNVYKONPKRSPFSEND	1032
1552	Db	DAANVOQLMEVRNLLGLGNA---GNDNADGNQVNIADIKKDPNSGSGSN	1597

```
RESULT 9
US-11-013-759-13
; Sequence 13, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-921MIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1992
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-13

Query Match      2.2%; Score 145; DB 7; Length 1992;
Best Local Similarity 17.6%; Pred. No. 0.054;
Matches 220; Conservative 161; Mismatches 430; Indels 438; Gaps 55;

Qy 44 VNSPELNHNKHTIYDSYEDVNNKLNFSFVENKSVKKRSLSPINNKTSDYIIPPSYS 103
Db 527 VTAFTYNGVKTTELNSDGTDFKESVKGSGTNNSLVTAHLASVLYNEVNRADSALQSFT 586

Qy 104 YR---NDKFNLSNEDNSGNTNNEFANTSEISI-----GKDNKQYTFIQRTHLFACGI 156
Db 587 VKEEDDDANAITVAKOTTKNAGA-----VSLKLGKNG--LIVATKOGTGVTFGL 636

Qy 157 KRKIKWICRE--NSEKIVCVPRKTIQLCVANFLNRLRLETMEKFEIFILSVNTEAKLL 214
Db 637 SQDGLTIGKSTLNDGLTVKDTNEQIQV--GANGI-----KFTNV----- 675

Qy 215 YNKNEGKDPIS-FCNELNRSFDSFRSFIQDDMPGNTDRVKGVIYNTKFSYYKEKVE 273
Db 676 -----NGSNPGTGIANTR-----ITRDKIGFAGS-----DGAVDT----- 706

Qy 274 KLANIKKEWKEKNKANLWNHMI VNHKGNISKECAIIPAEBPQINLWIKENWENFLMEKKR 333
Db 707 -----NKPYLDQDKLVGNVKITWTGNA----- 730

Qy 334 LFLNIKDKCVENKYEACFGGCRPLPCSSYT---SPMKSKTQMEVLTNLY-KKQNSGVVDKN 390
Db 731 -----GKAITGLSPTLPSTADQSSRNIELGNTIQDKRKSNAASIN 771

Qy 391 NFLNDLFKNNKNDLDDFFKKEKEVDLDCRCYATATIKSLPLNGPAKNDVDIASQINVD 450
Db 772 DILNTGNLKNNAPIDFVSTYDIDVDFANGNATTATVTHDTANTSK-----VVYDVNVD 827

Qy 451 LRGFPCNYSNNKSNWNTCTGFTTNKFPCTCEPPRRQTLCLGRT---YLLHRGHEE----- 502
Db 828 -----TTIHLTGDDNKKLGKGVTKTLNKTSAANGNTATNFVNSSDEDALVNA 874

Qy 503 -DYKEHL--LGASITY-----EAQLKYKYEKENALCSIIQNSYADLADIKGSIIK 553
Db 875 KQIAENLTLAKEIHTTKGTADTALQTFVKKVDE-----NNNADANAIT----- 920

Qy 554 DYVGKMEENLNKYNKDKKRNEESIKIFREK-----WWDENKENYKWSAVL-----K 602
Db 921 --VGQKUAN--NQWTLTLKGENGLNITKTDKNGVTFTGINTTSLGKAGKATLNDGGLSIK 976

Qy 603 NK-----ETCKDYDKFQKI----- 1035
Db 977 NPTSEQIQVADGVKFAKVNNGVVGAGIDGTTITRDEIGFTGTNGSLDKSKPHLSK- 1035

Qy 623 FKWGGDDFCERKEKEKIYSFESPKVECKKDC-----DENT-CKNKCSEYKKWIDLKKS 674
Db 1036 -----DGINAGGKKTITNIQSGBIAQNASHDAVTTGGKIYDLKTELENKISISAKTQNSLH 1089

Qy 675 EY---EKQVDKVTDKNNKMYD-----NIDEVK----- 699
Db 1090 EFSVADEQGNFTVSNPYSSYDTSKTSVDITPAGENGITTKVNVGVVRVCIQDTKGITTP 1149

Qy 700 -----NKEANYLKEKSKCKEKNVDFDDKI-- 723
Db 1150 KLTGVNNGKGVIVDSQNGQNTITGLSNTLANVTNDKGSVRTTEQGNIIKD---EDKTRA 1206

Qy 724 -----FNESPN-EYEDMCKKCEIKYLN-----EIKY---PKTKHDIYDIDTFS 763
Db 1207 ASIVDVLVSAGFNLQNGEAVDFVSTYDTVNFADGNATTAKVTVYDDTSTSKSVVYDVNV-- 1264

Qy 764 DTFGDGTPISI-----NANINEQQSGKDTSTNGSETSDSPVSHESPSDAAINVVKL 815
Db 1265 -----DDTTEVKDKKLGKVTTLTSTGTGANKFALSNOATGDALYK---ASDIVAHLNTL 1317

Qy 816 SGD-----ESSSETRGILDINDPSV--TNNVNEVHDASNTQGSVSNSTSDI----- 858
Db 1318 SGDIQAKGASQANNGAGYVDADGNKVIYDSTDNKYVQAKN-DGTVDKTEKVAKDLVAQ 1376

Qy 859 --TNGHSESLNRTTNAQDIKIGRSGNEQSDNQENS-----SHSSDN---SGSLTIGOV 907
Db 1377 AQTPTDGLTQAMNYKSVINKEQVNDANKQGINEDNAPVKGLEKAAADNKTNAAVTVGDL 1436

Qy 908 PSBDTQNTVDSQNPHRDTPNALASLP-----SDDKINETEGPDS----- 947
Db 1437 NAVAQPTLPFAG-----DTGTTAKLGETLTIKGGQDITNKLTDNNIGVAGTDTGFTVKL 1491

Qy 948 SRDSEN-----GRGDTTNTSHDVRTNIVSERRVNSHDFIRNGM-----AN 988
Db 1492 AKOLTNLNSVAGGTKIDDKGVSVFVDSGQAKANTPVLSSANGLDLGGKVISNVGKGTDT 1551

Qy 989 NNAHQYITQIEN-----NGIIRGQESAGNSVYKNDPKRSNFSSEND 1032
Db 1552 DAANVQQLNEVRNLLGLGNA---GNDNADGNQVNIADIKKDPNPGSSSN 1597

RESULT 10
US-11-013-759-4
; Sequence 4, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-921MIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2047
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-4

Query Match      2.2%; Score 145; DB 7; Length 2047;
Best Local Similarity 17.6%; Pred. No. 0.056;
Matches 220; Conservative 161; Mismatches 430; Indels 438; Gaps 55;

Qy 44 VNSPELNHNKHTIYDSYEDVNNKLNFSFVENKSVKKRSLSPINNKTSDYIIPPSYS 103
Db 582 VTAFTYNGVKTTELNSDGTDFKESVKGSGTNNSLVTAHLASVLYNEVNRADSALQSFT 641
```



```
Db 976 --VGQKXAN--NQVNTLTKGKGLNLTQDKNGVTGFGINTSLGKAGKSTLNDGGLSIK 1031
Qy 603 NK-----ETCKDYKFKQKI-----
Db 1032 NPTGSEQIQVGADGVKFAKVNNGNVGAGIDGTTTRITRDEIGFTGTGSLDKSKPHLSK- 1090
Qy 623 FKWDGDFCEKRKEKIVSFESFKVECKKDC-----DENT-CKNKCEYKKWIDLKKS 674
Db 1091 -----DGINAGGKKITINOSGEIAQNSHDAVTGKGIYDLKTELENKISSAKTAQNSLH 1144
Qy 675 EY---EKQVDKYTKDKNKKMYD-----NIDEVK----- 699
Db 1145 EFSVADEQGNFTVSNPYSDTSKTSVDTTFAGENGIITKVNKGVRVIGDQTKGLTTP 1204
Qy 700 -----NKEANYLKEKSKCKDVNPFDDKI-- 723
Db 1205 KLTVGNNNGKVIDSQNGQNTITGLSNTLANVTNDKGSVRTTEQGNIIKD---EDKTRA 1261
Qy 724 -----FNSPN-EYEDMCKKCKDEIKYLN-----EIKY---PKTKHDIYDIDTFS 763
Db 1262 ASIVDVLSAGFNQNGGEAVDFVSTYDTNPFADGNATTAKVTYDDTSTSKVYVDNV-- 1319
Qy 764 DTFGDGTPISI-----NANINEQSGKDTSTNGNSETSDSPVSHPEPSDAAINVEKL 815
Db 1320 -----DDTIEVKKLGKVTTLTSTGTGANKFALSQAQTDALVK---ASDIVAHLNLT 1372
Qy 816 SGD-----ESSSETRGILINDPSV--TNVNVNEVHDASNTQGSVNTSDI----- 858
Db 1373 SGDIQTAKGASQANSAGYVADGDKVYIDSTDNKYQAKN-DGTVDKTKVEAKDKLVAQ 1431
Qy 859 --TNGHSESSLNRTNAQDIIKGRSGNEQSDNQENS-----SHSSDN-----SGSLTIGQV 907
Db 1432 AOTPDGTLAQMNVKSVINKEQVNDANKQGINEDNAFVKGLEKAASDNKTKNAAVTVGDL 1491
Qy 908 PSEDNTQNTYDSQNPDRTPNALASLP-----SDDKINEIEGDFS----- 947
Db 1492 NAVAQTELTAG-----DTGTTAKUGLGETITIKGGQDQTNKLTNNNGVAGTDGFTVKL 1546
Qy 948 SRDSEN---GRGDTTNTDVRRTNIVSRRVNSHDFIRNGM-----AN 988
Db 1547 AKOLITLNSVNAAGTKIDDKGVSFVDSGGQAKANTPVLISANGLLDGGKVISNVGKGTDT 1606
Qy 989 NNAHQVITQIEN-----NGLIIRQESAGNSVNYKONPKRNPSSND 1032
Db 1607 DAANVOQLNEVRNLLGLGNA---GNDNADGNQVNIADIKKDPNSGSSN 1652

RESULT 12
US-11-077-550-100
; Sequence 100, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR FILING DATE: 1999-02-23
; PRIOR FILING DATE: 1997-08-22
; PRIOR FILING DATE: 1996-12-27
```

```
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 100
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-100
```

```
Query Match 2.2%; Score 141; DB 7; Length 865;
Best Local Similarity 18.7%; Pred. No. 0.036;
Matches 181; Conservative 138; Mismatches 344; Indels 306; Gaps 45;

Qy 252 TDRVKGVIYTKFSDYYKEKNVEKLNLIK---EWEKKNKANLWNHMI VNHKGNISKECA 307
Db 46 TDRI-WILPERYTFGYPEDFNKSSGIFNRDVCVEYDPPYLN-----TNDKKNI----- 93
Qy 308 IIPAEFPQINLWIKENENFLMEKKRLFLNIKDKCVENKKYACFGGC-----RLPCSS 361
Db 94 -----FLOTMIKLENRIKSKPLGKLEMLINGIPYLGDRRVPLEE 134
Qy 362 YTSFMKSKTQMEVLNLYKKNSGVDKNLFNLDLFKNKNNDLDDFFKKEKEYDDLCDC 421
Db 135 FN-----TNI-----ASVTNKLISNPGEVERKGI--FAN----- 163
Qy 422 RYTATTIIKSFLANGPAKND--VDIASQINVDNLGRFG-----C-NYKS--NNEKSMNC 468
Db 164 -----LIIFGPGVULNENETIDIGIHFASREGFGGIMQMFKCPBYVSVFNVQB--- 214
Qy 469 TGTFTNKFPGTCPPRRQTLCLGRTYLLHRGHEEDYKEHLGLGASIVBAQILLKYKKEKDE 528
Db 215 ----NK-----GASIFNRGRGYSDPALILMHHLIHLVHGL-YGIKVDD- 252
Qy 529 NALCSIIQNS---YADLADIKGSII-----KDYCKKME-----ENL 564
Db 253 ---LPIVPNEKFFMQSTDAIQAEELYTFGGQDPSIITPSTDKSIYDKVLQNFGRGIVDRL 309
Qy 565 NYKNDKKNNEESLKIIFREKWD-----ENKENVWVMSAVLNKKECTCKDYDFQKIPQF 619
Db 310 NKVLVCISDPNININIKKFKDKYFVEDSEKYSI-----DVESFDKLYKS 357
Qy 620 LRW-FKEW--GDDFCERKREKIYSPESFKVECKKDCDENTCKNKCEYKKWIDLKKEY 676
Db 358 LMFQFTETNIAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYVIEEGFNISDKDM 411
Qy 677 EKQVDKYTKDKNKKMYDNDIDEVKNKEANY-----LKEKSKCKDVNFDKIFNESPN 729
Db 412 EKEYRGQNKAINKQAY---BEISKEHLAVYKIQMCGLVPGRSGPCIDVDNEDLFFIADKN 468
Qy 730 EYEDMCKKCKDEIKYLNKIKYPTKHDYDIDTSDTFGDCGTPISINANINEQSGKDTSN 789
Db 469 SFSDDLKSKNERIEYNTQSVNIENDFPINELILDTDL-----ISKIELPRENTE 516
Qy 790 TGNSETSDSPVSHPEPSDAAINVEKLSGDESS-----SETRGILINDPSVNNVNEVH 843
Db 517 SLTDFNVDPVVEKQPA-----IKKIFTDENTIFQVLYSQTFP-LDIRDISLTSSFPDAL 570
Qy 844 DASNTQGSVNTSDIITNGHSESSLNRTTNAQDIKGRSG-----NEQSDN-----QENSSH 894
Db 571 LFSNKVYSPFSDMYI-----KTAN---KVEAGLFAGVKQIVNDVFIEANKSN 616
Qy 895 SSDNSG-----SLTIGQVPSDENTONTYDSQN-----PHRTPNALASLPS 935
Db 617 TMDKIADISLIVPIYIGLALNVGNETAKGNFENAFETAGASILLEFPELLIPVVGAFLE 676
Qy 936 ---DDKINEIEGFDSSRDSENGRGD-----TTSNT--HDRV----- 966
Db 677 SYIDNKNKIKTIDNALTKRNEKWSDMYGLIVAQLSTVNTQFTIKEGMVKALNYQAQA 736
Qy 967 -----RTNIVSERRVNSHDFIRNGMAN--NNAHQVITQIEN--NG----- 1003
```

Db 737 LEBIIKRYNIYSEKESKNINIDFNDINSKLNQINQAINNFINNGSVSYLMKMP 796  
 Qy 1004 -----IIRGOESAGNSVNYKDNPKRNSFSENHKKNIQEYNSR-----DTKRVREI 1054  
 Db 797 LAVELKDFDNTLKKNLNLYIDENKLYLIGSAEYKSVNKLTKIMPFDLSIYNTDIL 856  
 Qy 1055 LSKONKCNN 1063  
 Db 857 IEMFNKNS 865

## RESULT 13

US-11-077-550-80  
 ; Sequence 80, Application US/11077550  
 ; Publication No. US20050244435A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shone, Clifford Charles  
 ; APPLICANT: Quinn, Conrad Padraig  
 ; APPLICANT: Foster, Keith Alan  
 ; APPLICANT: Chaddock, John  
 ; APPLICANT: Marks, Philip  
 ; APPLICANT: Sutton, J. Mark  
 ; APPLICANT: Stancombe, Patrick  
 ; APPLICANT: Wayne, Jonathan  
 ; TITLE OF INVENTION: Recombinant Toxin Fragments  
 ; FILE REFERENCE: 1581.0130004  
 ; CURRENT APPLICATION NUMBER: US/11/077,550  
 ; CURRENT FILING DATE: 2005-03-11  
 ; PRIOR FILING DATE: 2002-09-12  
 ; PRIOR FILING DATE: 2002-09-12  
 ; PRIOR FILING DATE: 1999-02-23  
 ; PRIOR FILING DATE: 1997-08-22  
 ; PRIOR FILING DATE: 1997-08-22  
 ; PRIOR FILING DATE: 1996-12-27  
 ; PRIOR FILING DATE: 1996-12-13  
 ; PRIOR FILING DATE: 1996-08-23  
 ; NUMBER OF SEQ ID NOS: 179  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 80  
 ; LENGTH: 867  
 ; TYPE: PRP  
 ; ORGANISM: Clostridium botulinum  
 US-11-077-550-80

Query Match 2.2%; Score 141; DB 7; Length 867;  
 Best Local Similarity 18.6%; Pred. No. 0.036;  
 Matches 181; Conservative 138; Mismatches 344; Indels 308; Gaps 45;  
 Qy 252 TDRVGVINTKFSYDYKKNVKNLNIKK-----EWEKKNANLNMHNVHKNISKECA 307  
 Db 46 TDRI-WIIPERYTGYPFNKSGIFNRDVCYDYLN-----TNDKKNI----- 93  
 Qy 308 IIPAEPPQINLIWKENWENFLMEKKRLFLNIDKCVENKYVACFGGC-----RLPCSS 361  
 Db 94 -----PLQWIKLFLNRKSKPLGKLEMIINGIPYLGDRRPLEE 134  
 Qy 362 YTSFMKSKTQMEVLTNLYKKNSGVKNPNFLNLPKNNKNDLDDPFKEKEYDDLCDC 421  
 Db 135 FN-----TWI-----ASVTNKLINPGEVERKKGI-----PAN----- 163  
 Qy 422 RYTATIISKFLNGPAKND---VDIASQINVDLGRFG-----C-NYKS---NNEKSWNC 468  
 Db 164 -----LIIFGPGVLNENETIDIGIHNHFAASREGFGIMQMKFCPEYVSFVNNVQE--- 214  
 Qy 469 TGFTNKPFGTCBPRTQCLGRTYLLHRGHEDEYKHLGASIEYEAQLLKYKKEDE 528  
 Db 215 -----NK-----GASIFNRGRGYSFDPALILMHHLIHLGL-YGIKVDD- 252

Qy 529 NALCSIIIONS---YADLADIIGSDII-----XDYYGKOME-----ENL 564  
 Db 253 ---LPVPNEKKYFMQSTDAIQABELYTFGQDPSIITPSTDKSIYDKVLQNFPGIVDRL 309  
 Qy 565 NKVNKDKRNEESLKIPREKWD-----ENKENVWKMVAVLKVKETCKDYKFKQIPQF 619  
 Db 310 NKVLVCISDPNININIKYKFKDYKFEVDESEKYSI-----DVESEPKLYKS 357  
 Qy 620 LRW--FKEW--GDDPCEKREKIYFESFKVECKKDCDENTCKNKCSEYKWKWIDLKSEY 676  
 Db 358 LMGFTETNAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYIEBGFNISDKDM 411  
 Qy 677 EKQVDKYTKDKNKMVDNIDDEVKNKEANVYLKEKSE-----CKDNFDDKIFNES 727  
 Db 412 EKEYRGONKAINQAY---EESKEHLAVYKIQWCKSDDDKAPGICIDVDNEDLFFIAD 468  
 Qy 728 PNEYEDCKKDEIKYLYNEIKYPTKHDIDYDITFSTFGDGTPIISINAINNEQOSKDT 787  
 Db 469 KNSFSDLSKNERIEYNTQSYIENDPFINELILDIDL-----ISKIELPREN 516  
 Qy 788 SNTGNSSETSPVSHPESDAAINVEKLSGDESS-----SETRGILIDINDPSVTNNVNE 841  
 Db 517 TESLTDNFVDVPVEKQPA-----IKKIFTDENITFOYLSQTFP-LDIRDISLTSFDD 570  
 Qy 842 VHDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSG-----NEQSDN-----QENS 892  
 Db 571 ALLFSNKVYFFSMDYI-----KTAN-----KVVEAGLPAGWVKQIVNDFVIEANK 616  
 Qy 893 SHSDNSG-----SLTIGQVPSDONTONTYDSQN-----PHRDTNALASL 933  
 Db 617 SNTMDKIADISLIVPYTIGLALNVGNETAKGNFENAFIAGASILLLEPIPELLIPVVGAF 676  
 Qy 934 PS---DDKINEIRGFSRSENGRGD-----TTSNT---HDOVR----- 966  
 Db 677 LESYIDNKKIKITIDNALTNRNEKSDMTGLIVAQWLSVNTQFYTIKEGMYKALNYQA 736  
 Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHQVITQIEN--NG----- 1003  
 Db 737 QALEEIIKRYNIYSEKESKNINIDFNDINSKLNQINQAINNFINNGSVSYLMKMP 796  
 Qy 1004 -----IIRGOESAGNSVNYKDNPKRNSFSENHKKNIQEYNSR-----DTKRVREI 1052  
 Db 797 IPLAVEKLDFTNLTKNLNLNLYIDENKLYLIGSAEYKSVNKLTKIMPFDLSIYNTD 856  
 Qy 1053 ILSKONKCNN 1063  
 Db 857 ILEMFNKNS 867

## RESULT 14

US-11-077-550-22  
 ; Sequence 22, Application US/11077550  
 ; Publication No. US20050244435A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shone, Clifford Charles  
 ; APPLICANT: Quinn, Conrad Padraig  
 ; APPLICANT: Foster, Keith Alan  
 ; APPLICANT: Chaddock, John  
 ; APPLICANT: Marks, Philip  
 ; APPLICANT: Sutton, J. Mark  
 ; APPLICANT: Stancombe, Patrick  
 ; APPLICANT: Wayne, Jonathan  
 ; TITLE OF INVENTION: Recombinant Toxin Fragments  
 ; FILE REFERENCE: 1581.0130004  
 ; CURRENT APPLICATION NUMBER: US/11/077,550  
 ; CURRENT FILING DATE: 2005-03-11  
 ; PRIOR FILING DATE: 2002-09-12  
 ; PRIOR FILING DATE: 2002-09-12  
 ; PRIOR FILING DATE: 1999-02-23  
 ; PRIOR FILING DATE: 1997-08-22  
 ; PRIOR FILING DATE: 1997-08-22



Qy 529 NALCSIIQNS---YADLADIIGSDII-----KDYVGKOME-----ENL 564  
Db 253 ---LPIVNEKFFWQSTDAIQABELYTFGQDPSIITPSTDKSIYDKVLQNFQGVDR 309  
Qy 565 NKVNKDKRNEESLUKIPREKWD-----ENKENVWKMVAVLKNKTCCKDYKFKQKIPQF 619  
Db 310 NKVLVCISDPNININIKNFKDKYKFVEDSEGKYSI-----DVESFDKLYKS 357  
Qy 620 LRW--PKW---GDDPCERKEKIYSPESFKVECKKODCENTCKNKCSEYKWIIDLKSEY 676  
Db 358 LMFQFTETNAENYKIKTRASYFSDSLPPVKI-----NLLDNEIYTBEGFNISDKM 411  
Qy 677 EKQVDKYTKDNKKMYDNID-----EVKNKEANVYLKE-----KSKCEKDVNFD 721  
Db 412 EKEYRGONKAINKOAVEISKEHLAVYKIOMCKSEKLYDDDDKORWGSSRICIDVND 471  
Qy 722 KIFNESPEYEDMCKCDEIKYLNEIKYPTKHDYIDIDTFSDTFGDTGPISINANEQ 781  
Db 472 LFTADKNSFSDLSKNERIEYNTQSNYIENDPFINELILDIDL-----LSKI 519  
Qy 782 QSGKOTNTGNSSETSPVSHPESDAAINVEKLSGDESS-----SETGILDINPVS 835  
Db 520 ELPRENTESLTFNVDVPVYEKQPA-----IKKIFTDENTIFQYLYSQTTP-LDIRDISL 573  
Qy 836 TNNVNEVHDASNTQGSVNTSDITNGHSESSLRTTNAQDIKGRSG-----NEQSDN-- 888  
Db 574 TSSFDLALFSNKVYSPFSMDYI-----KTAN-----KVVEAGLFGWVKQIVNDF 619  
Qy 889 --GENSSHSDNSG-----SLTIGQVPSDNTQNTYDSON-----PHRDTP 927  
Db 620 VIEANKNTWDKTDADISLIVPYIGLALNVGNETAKGNFENAFETAGASILLFPELLIP 679  
Qy 928 NALASLPS---DDKINEIEGDFSRSRSENGRGD-----TTSENT--HDVR----- 966  
Db 680 VVGAFLLESYIDNKKIITIDNALTKRNEKWSMDMYGLIVAQWLSTVNTQFYTIKEGMYK 739

## RESULT 16

US-11-077-550-82  
; Sequence 82, Application US/11077550  
; Publication No. US2005024435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22

; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 82  
; LENGTH: 876  
; TYPE: PRT  
; ORGANISM: Clostridium botulinum  
; US-11-077-550-82

Query Match 2.1%; Score 138.5; DB 7; Length 876;  
Best Local Similarity 18.5%; Pred. No. 0.051;  
Matches 181; Conservative 140; Mismatches 345; Indels 311; Gaps 45;

Qy 252 TDRVKGVIINTKFSYKKEVKEKLNIIKK-----EMWEKKNKANLWNHIVNKHGKISCECA 307  
Db 46 TDRI-WIIPERYTFGYPEDFNKSSGIFNRDVCYDDPDYLN-----TNDKKN----- 93  
Qy 308 IIPAEBPQINLWIKENWENFMEKQRULFLNKKCKENKKEVACFGG-----RLPCSS 361  
Db 94 -----FLQTMIKLFNRKSKPLGKLEMIINGIPVYLGDRRVPLEE 134  
Qy 362 YTSFMKSKTKQMEVLNLYKKKSGVDKNNFLNDLFKKNKNDLDDFFKKEKEYDLDLDCD 421  
Db 135 FN-----TNI-----ASVTNKLISNPGSEVERKGGI-----FAN----- 163  
Qy 422 RYTATIISKFLNGPAKND---VDIASQINVDNLRGFG-----C-NYKS---NNEKSWNC 468  
Db 164 -----LIIFGPGVLNENETIDIGIHFASREGFGGIMQMKFCEPVSVFNNVQB--- 214  
Qy 469 TGTFTNKPCTCPRRQTLCLQRTYLLHRGHEDYKEHLGLGASIEYEAQLLKYYKKKDE 528  
Db 215 ---NK-----GASIFNRGRGYFSDPALILMHHLIHLVHGL-YGIKYVDD- 252  
Qy 529 NALCSIIQNS---YADLADIIGSDII-----KDYVGKOME-----ENL 564  
Db 253 ---LPIVNEKFFWQSTDAIQABELYTFGQDPSIITPSTDKSIYDKVLQNFQGVDR 309  
Qy 565 NKVNKDKRNEESLUKIPREKWD-----ENKENVWKMVAVLKNKTCCKDYKFKQKIPQF 619  
Db 310 NKVLVCISDPNININIKNFKDKYKFVEDSEGKYSI-----DVESFDKLYKS 357  
Qy 620 LRW--PKW---GDDPCERKEKIYSPESFKVECKKODCENTCKNKCSEYKWIIDLKSEY 676  
Db 358 LMFQFTETNAENYKIKTRASYFSDSLPPVKI-----NLLDNEIYTBEGFNISDKM 411  
Qy 677 EKQVDKYTKDNKKMYDNID-----EVKNKEANVYLKE-----KSKCEKDVNFD 721  
Db 412 EKEYRGONKAINKOAVEISKEHLAVYKIOMCKSEKLYDDDDKORWGSSRICIDVND 471  
Qy 722 KIFNESPEYEDMCKCDEIKYLNEIKYPTKHDYIDIDTFSDTFGDTGPISINANEQ 781  
Db 472 LFTADKNSFSDLSKNERIEYNTQSNYIENDPFINELILDIDL-----LSKI 519  
Qy 782 QSGKOTNTGNSSETSPVSHPESDAAINVEKLSGDESS-----SETGILDINPVS 835  
Db 520 ELPRENTESLTFNVDVPVYEKQPA-----IKKIFTDENTIFQYLYSQTTP-LDIRDISL 573  
Qy 836 TNNVNEVHDASNTQGSVNTSDITNGHSESSLRTTNAQDIKGRSG-----NEQSDN-- 888  
Db 574 TSSFDLALFSNKVYSPFSMDYI-----KTAN-----KVVEAGLFGWVKQIVNDF 619  
Qy 889 --GENSSHSDNSG-----SLTIGQVPSDNTQNTYDSON-----PHRDTP 927  
Db 620 VIEANKNTWDKTDADISLIVPYIGLALNVGNETAKGNFENAFETAGASILLFPELLIP 679  
Qy 928 NALASLPS---DDKINEIEGDFSRSRSENGRGD-----TTSENT--HDVR----- 966  
Db 680 VVGAFLLESYIDNKKIITIDNALTKRNEKWSMDMYGLIVAQWLSTVNTQFYTIKEGMYK 739

```

Qy 967 -----RTNIVSERRVNSHDFIRNGMAN--NNAHHQYITQIEN--NG----- 1003
Db 740 ALNYQAALBEEIKYRYNIYSEKESNINIDFNIDNSKLNQINQAINNNFNGCSVS 799
Qy 1004 -----IIRQESAGSNVYKNDPKESNFSSENDHKKNIQEYNSR-----DTK 1046
Db 800 YLMKKMPLAVEKLLDFDNTLKNLLNYIDENKLYLIGSAEYKSKVNYKYLKTMPPDLS 859
Qy 1047 RVREIILKSKONKCNN 1063
Db 860 IYTNDTILIEMNKYN 876

RESULT 17
US-11-077-550-106
; Sequence 106, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 03/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 106
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-106

```

```

Query Match      2.1% Score 136.5; DB 7; Length 876;
Best Local Similarity 18.5%; Pred. No. 0.066;
Matches 181; Conservative 139; Mismatches 346; Indels 311; Gaps 45;

Qy 252 TDRVKGINTKFSYKQVKEKUNNKK-----BWEKKNANLWNHNVHKNISKECA 307
Db 46 TDRI-WIPERYTFGYPKDFNKSNGIFNRDVCYYDPYLN-----TNDKKN----- 93
Qy 308 IIPAEPPQINLWIKENWENFLMEKKRLPLNKKCKVENKKEACFGCC-----RLPCSS 361
Db 94 -----FLOTWIKLFNRKSKPLGKLELMIINGIPYLGORRVPLEE 134
Qy 362 YTFPMKSKTQMEVLTNLYKKNSGVDRNNFLNDFPKNNKNDLDDFFKNEKEYDDLCDC 421
Db 135 FN-----TNI-----ASVTNKLISNPGVERKGI-----FAN----- 163
Qy 422 RYATATIKSFLNGPAKD-----VDIASQINVDNLRGFG-----C-NYKS--NNEKSWNC 468
Db 164 -----LIIFGPGVLTNENETIDIGIQNHFASREGFGGIMQMKPCPEYVSFVNNVQV--- 214
Qy 469 TGTFTNKPFGTCPEPRRQTLCLGRTYLLHRGHEEDYKEHLIGASIEYEAQLLKYYKEKDE 528

```

```

Db 215 -----NK-----GASIFNRRGYFSDPALIMHLELIHVLHGL-YGIKVDD- 252
Qy 529 NALCSTIONS---YADLADIIGSDII-----KDYTGKME-----ENL 564
Db 253 ---LPVNPKEKFFMQSTDAIQAEELYTFGGQDPSIITPSTDKSIYDKVLQNRGIVDRL 309
Qy 565 NKVNDKKKNESLKLIFREKWD-----ENKENVWKMGAVALKNKCKDYDFQKIPQF 619
Db 310 NKVLVCISDPNINITYKNFKDKYKRFVEDSEKYSI-----DVESFDKLYKS 357
Qy 620 LRW-FKEW--GDDFCERKEKIIYSPFSFKVECKKDDCDEKCNKCEYKKWIDLKSEY 676
Db 358 LMFGFTETNIAENYKIKTRASYFSDSLPVKIK-----NLLDNEIYTIIEGFINISDKM 411
Qy 677 EKQVDKYTKDKNKKMYDNID-----EVKNKEANVYLKE-----KSECKEDVNFDD 721
Db 412 EKEVRGONKAINQAAYEEISKEHLAVYKIQMCKSEKLYDDDDKDRWSSRICIDVDNED 471
Qy 722 KIFNESPNEYEDMCKKDEIKYLNKIKYPTKHDIYDIDTFSDTFGGGTPIISANINEQ 781
Db 472 LFFIADKNSFSDLSKNERIEYNTQSNYIENDPFINELILDTDL-----ISKI 519
Qy 782 QSGKDTSGTSGSETSDSPVSHPEPESDAAINVEKLSGDESS-----SETRGILDINDPSV 835
Db 520 ELPESTESLTDNFVDVPVTEKQPA-----IKKIFTDENTIFYLYSQTTPP-LDIRDISL 573
Qy 836 TNNVNEVHDASNTQGSVSNSTSDITNGHSBSSLNRTTNAQDIKICRSQ-----NEQSDN-- 888
Db 574 TSSFDALLFSNKVYFFSMDYI-----KTAN-----KVEAGLFAGWVKQIVNDF 619
Qy 889 --QENSSHSDNSG-----SLTIGQVPSDONTQNTYDSQN-----PHRDTP 927
Db 620 VIEANKSTMDKIADISLIVPYTIGLALNVGNETAKGNFENAFIAGASILLEFIPELLIP 679
Qy 928 NALASLPS---DDKINEIGFSDSRDSENGRGD-----TTSNT--HDVR----- 966
Db 680 VVGAFLLESYIDNKKIITIDNALTNRNEKWSDMYGLIVAQLSVNTVNTQFYTIKEGMYK 739
Qy 967 -----RTNIVSERRVNSHDFIRNGMAN--NNAHHQYITQIEN--NG----- 1003
Db 740 ALNYQAALBEEIKYRYNIYSEKESNINIDFNIDNSKLNQINQAINNNFNGCSVS 799
Qy 1004 -----IIRQESAGSNVYKNDPKESNFSSENDHKKNIQEYNSR-----DTK 1046
Db 800 YLMKKMPLAVEKLLDFDNTLKNLLNYIDENKLYLIGSAEYKSKVNYKYLKTMPPDLS 859
Qy 1047 RVREIILKSKONKCNN 1063
Db 860 IYTNDTILIEMNKYN 876

```

```

RESULT 18
US-11-077-550-108
; Sequence 108, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 03/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273

```



```
Db 215 ----NK-----GASIFNRGYSFSDPALILMHLEHIVHGL-YGIVKDD- 252
Qy 529 NALCSIIIONS---YADLADIIGSDII-----KDYGKOME-----ENL 564
Db 253 ---LPIVNEKFFMQSTDAIQABELYTFGQDPSIITPSTDKSIYDKVLQNRGIVDRL 309
Qy 565 NKYNKDKRNEESLKIFREKWD-----ENKENVWVMSAVLKNKTKCDYDFQKIPQF 619
Db 310 NKVLVCISDPNININIKYKPKDYKVFEDSEKYSI-----DVSFDFKLYKS 357
Qy 620 LRW-FKEW---GDDPCERKEKIYSFESFKVECKKDCDENTCKNKCSEYKWKIDLKSEY 676
Db 358 LMFGETETNAENYKIKTRASYSFDSLPPVKIK-----NLLDNEIYTIIEGPNISDKM 411
Qy 677 EKQVDKYTKDKNKKMYDNIDDEVKNKEANVYLKESKE-----CKDVNFDKIFNESP 728
Db 412 EKEYRGONKAINQAY---EESKEHLAVYKIQMCSAIEGRAPGICIDVDNEDLFFIADK 468
Qy 729 NEYEDMCKKDEIKYLNKPKTKHDYIDIDFSDTFGDTFGDTPISINANINEQSGKDT 788
Db 469 NSPDDLSKNERIEYNTQSNYIENDFPINELILDIDL-----ISKIELPSENT 516
Qy 789 NTGNSSETSDSPVSHPEPSDAAINVEKLSGDSESS-----SETRGILDINDPSVTNNVNEV 842
Db 517 ESULTDFNVDPVVEKQPA-----IKKIFTDENTIFQYLYSQTFP-LDIRDISLTSFDDA 570
Qy 843 HDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSG-----NQSDN-----QENSS 893
Db 571 LLFSNKVYSPFSDYI-----KTAN---KVEAGLFAGVWKQIVNDFVIEANKS 616
Qy 894 HSSDNGS-----SLTIGQVPSEDTQNTYDSQN-----PHRDTNPALASLP 934
Db 617 NTMDKIADISLIVPYIGLALNVGNETAAGNFENAFETAGASILLEFTEPELLIPVVGAFLL 676
Qy 935 S---DDKINEIEGFDSSRDSSENGRGD-----TTSNT---HDVR----- 966
Db 577 ESYIDNKNKIKITDIALTRNEKWSMDYGLIYVQWLSTVNTQPTTIKEGMYKALYQAO 736
Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHHQYITOEN--NG----- 1003
Db 737 ALBEEIKYRYNIYSEKESKININIDFNSKLNBEINQAINNINNFINGCSVSLMKKMI 796
Qy 1004 -----IIRQOESAGNSVNYKDKPKRSNFSSENHDKNQIENSR-----DTRKVRBEII 1053
Db 797 PLAVEKLLDFNTLKNLNYIDENKLYLIGSAEYKSKVNYKYLKTIIMPFDLSIYNTDTI 856
Qy 1054 KLSKQKNN 1063
Db 857 LIEMFNKYS 866
```

```
RESULT 20
US-11-077-550-175
; Sequence 175, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stacombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
```

```
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 175
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-175
```

```
Query Match 2.1%; Score 133.5; DB 7; Length 860;
Best Local Similarity 18.6%; Pred. No. 0.097;
Matches 180; Conservative 137; Mismatches 346; Indels 307; Gaps 45;
```

```
Qy 252 TDRVGYINTKFSDDYKKNVKNLNNIKK-----EWMKKNANLWNNHIVNHKGNISKECA 307
Db 40 TDRI-WIIPERYTFGYKPEDFNKSGIFNRDVCYYDPOVLN-----TNDKKNI----- 87
Qy 308 IIPABEPQINLWIKENENFLMEKKELFLNLIKDKCVENKYEACFGC-----RLPCSS 361
Db 88 -----FLQTMIKLFNRKSKPLGKLELMIINGIPYLGDRRVPLEE 128
Qy 362 YTSFMKKSKTQMEVLNLYKKNSGVDKNNFLNDFKNNKNDLDDFFKNEKEYDDILCDC 421
Db 129 FN-----TNI-----ASTVVKLISNPGEVERKGI-----FAN----- 157
Qy 422 RYTATIIKSFLNGPAKND---VDIASQINVDNLGRFG-----C-NYKS--NNEKSWNC 468
Db 158 -----LIIFGPGVLNENETIDIGIQNHPASREGFGGIMQMFKFCPEYVSFVNNVQS--- 208
Qy 469 TGTFTNKPFCPTCEPPRQTLCLGRTYLLHRGHEEDYKEHLLGASLYEAOQLLYKYEKDE 528
Db 209 -----NK-----GASIFNRGYSFSDPALILMHLEHIVHGL-YGIVKDD- 246
Qy 529 NALCSIIIONS---YADLADIIGSDII-----KDYGKOME-----ENL 564
Db 247 ---LPIVNEKFFMQSTDAIQABELYTFGQDPSIITPSTDKSIYDKVLQNRGIVDRL 303
Qy 565 NKYNKDKRNEESLKIFREKWD-----ENKENVWVMSAVLKNKTKCDYDFQKIPQF 619
Db 304 NKVLVCISDPNININIKYKPKDYKVFEDSEKYSI-----DVSFDFKLYKS 351
Qy 620 LRW-FKEW---GDDPCERKEKIYSFESFKVECKKDCDENTCKNKCSEYKWKIDLKSEY 676
Db 352 LMFGETETNAENYKIKTRASYSFDSLPPVKIK-----NLLDNEIYTIIEGPNISDKM 405
Qy 677 EKQVDKYTKDKNKKMYDNIDDEVKNKEANVYLKESKE-----CKDVNFDKIFNESP 728
Db 406 EKEYRGONKAINQAY---EESKEHLAVYKIQMCSAIEGRAPGICIDVDNEDLFFIADK 462
Qy 729 NEYEDMCKKDEIKYLNKPKTKHDYIDIDFSDTFGDTFGDTPISINANINEQSGKDT 788
Db 463 NSPDDLSKNERIEYNTQSNYIENDFPINELILDIDL-----ISKIELPSENT 510
Qy 789 NTGNSSETSDSPVSHPEPSDAAINVEKLSGDSESS-----SETRGILDINDPSVTNNVNEV 842
Db 511 ESULTDFNVDPVVEKQPA-----IKKIFTDENTIFQYLYSQTFP-LDIRDISLTSFDDA 564
Qy 843 HDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSG-----NQSDN-----QENSS 893
Db 565 LLFSNKVYSPFSDYI-----KTAN---KVEAGLFAGVWKQIVNDFVIEANKS 610
Qy 894 HSSDNGS-----SLTIGQVPSEDTQNTYDSQN-----PHRDTNPALASLP 934
Db 611 NTMDKIADISLIVPYIGLALNVGNETAAGNFENAFETAGASILLEFTEPELLIPVVGAFLL 670
Qy 935 S---DDKINEIEGFDSSRDSSENGRGD-----TTSNT---HDVR----- 966
```

Db 671 ESYDINKNKKIITDNLTKRNEKWSMDYGLIVAQWLSVTNTQVYTIKBMGYKALNYQAQ 730  
 Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHQYITQIEN--NG-----1003  
 Db 731 ALBHIKRYNIYSEKESNINIDFNDINSKLNQINGINQAINNFINFGCSVYLMKMI 790  
 Qy 1004 -----IIRGOEBSAGNSVNYKQPNKRSNFSSENDHKHKNIOEYNSR-----DTRKVRBEII 1053  
 Db 791 PLAVEKLLDFNTLKNLLNIDENKLYLIGSABEYKSVKNYKLTIMPFDLSIYTNDTI 850  
 Qy 1054 KLSQKNCNN 1063  
 Db 851 LIEMFNKYS 860

## RESULT 21

US-11-077-550-94  
 ; Sequence 94, Application US/11077550  
 ; Publication No. US20050244435A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shone, Clifford Charles  
 ; APPLICANT: Quinn, Conrad Padraig  
 ; APPLICANT: Foster, Keith Alan  
 ; APPLICANT: Chaddock, John  
 ; APPLICANT: Marks, Philip  
 ; APPLICANT: Sutton, J. Mark  
 ; APPLICANT: Stancome, Patrick  
 ; APPLICANT: Wayne, Jonathan  
 ; TITLE OF INVENTION: Recombinant Toxin Fragments  
 ; FILE REFERENCE: 1581.0130004  
 ; CURRENT APPLICATION NUMBER: US/11/077,550  
 ; PRIOR FILING DATE: 2005-03-11  
 ; PRIOR FILING DATE: 2002-09-12  
 ; PRIOR APPLICATION NUMBER: 09/255,829  
 ; PRIOR FILING DATE: 1999-02-23  
 ; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
 ; PRIOR FILING DATE: 1997-08-22  
 ; PRIOR APPLICATION NUMBER: 08/782,893  
 ; PRIOR FILING DATE: 1996-12-27  
 ; PRIOR APPLICATION NUMBER: GB9625996.5  
 ; PRIOR FILING DATE: 1996-12-13  
 ; PRIOR APPLICATION NUMBER: GB9617671.4  
 ; PRIOR FILING DATE: 1996-08-23  
 ; NUMBER OF SEQ ID NOS: 179  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 94  
 ; LENGTH: 862  
 ; TYPE: PRT  
 ; ORGANISM: Clostridium botulinum  
 US-11-077-550-94

Query Match 2.1%; Score 133.5; DB 7; Length 862;  
 Best Local Similarity 18.6%; Pred. No. 0.097;  
 Matches 180; Conservative 137; Mismatches 346; Indels 307; Gaps 45;  
 Qy 252 TDRVKGYINTFSDYKKEKVEKLNKIKK-----EWEKKNALNWHMIVNHKHNISKECA 307  
 Db 42 TDRI-WIIPERYTFGYKPEDFNKSSGIFNRDVCYDYPDLN-----TNDKQNI----- 89  
 Qy 308 IIPABEPQINLWIKENWENFLWEKRLFLNLIKDKOVENKKEACFGG-----RLPCS 361  
 Db 90 -----FLOTWIKLFRNRSKPLGKLEMLINGIPYLGDRRVLPEE 130  
 Qy 362 YTFPMKSKTQMBVLTNLYKKNSGVDKNNFLNDFKKNKNKNDLDDFPKNEKEYDDLCDC 421  
 Db 131 FN-----TNI-----ASVTNKLISNPGVEKKG-----PAN----- 159  
 Qy 422 RYTAIYKISFLNGPAKD-----VDIASQINVDNLRGFG-----C-NYKS--NNEKSWNC 468  
 Db 160 -----LIIFGPGVFLNENETIDIGIHNHFASREGFGIMQMFKFCPEYVSFVNNVQE--- 210

Qy 469 TGFTNKFPGCTPPRRQTLGRYTLHRRGHEBDYKEHLIGASIVYBAQLIKYKKEKDE 528  
 Db 211 -----NK-----GASIFNRRGYFSDPALILMHHLIHLVGL-YGKVDD- 248  
 Qy 529 NALCSIIIONS---YADLADIKGSII-----KDYKQKME-----ENL 564  
 Db 249 ---LPIVPEKCKFMQSTDAIQABELYTFGQDPSIIITPSTDKSIYDKVLQNFGRGIVDRL 305  
 Qy 565 NKVNDKRRNEESLKIIFREKWD-----ENKENVWVMSAVLKNKTCCKDYKQKPOF 619  
 Db 306 NKVLVCLISDPNINININIKAFKDYKVFVESEKYSI-----DVEFDKLYKS 353  
 Qy 620 LRW-FKEW--GDDFCERKEKIYFSEFKVECKKCDENTCKNKCSEYKMWDLKKSEY 676  
 Db 354 LMFGETETNAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYTBEGFNISDKDM 407  
 Qy 677 EKQVDKYTKKKNKMYNDIDEVKNKEANVYLKESKE-----CKDVNPDDEKIFNESP 728  
 Db 408 EKEYRGQNKAINKOAY---EEISKEHLAVYKIQWCESAIEGRAPGICIDVDNEDLFFIADK 464  
 Qy 729 NEYEDMCKCDEIKYLNKIKYPTKHDIDIDTFSDTFDGTPTISINANINEQSGKDT 788  
 Db 465 NSFSDLSKNERIEYNTQSNYIENDFPINELILDTL-----ISKILPSENT 512  
 Qy 789 NTGNSSETSDSPVSHESPESDAAINVEKLSGDESS-----SETRGILINDPSVTNNVNEV 842  
 Db 513 ESLTDFNVDPVVEKQPA-----IKKIFTDENTIFQVLYSQTFP-LDIRDISLTSSPDDA 566  
 Qy 843 HDASNTQGSVNTSDITNGHSESLNRTTNAQDIKIGRSG-----NQSDN-----QENSS 893  
 Db 567 LLFSNKVYSPFSDYI-----KTAN---KVVYAGLIFAGWVKQIVNDFVIEANKS 612  
 Qy 894 HSSDNG-----SLTIGQVPSDNTONTYDSON-----PHRDTNALASLP 934  
 Db 613 NTMDKIADISLIVPIYIGLALNVGNETAKGNFENAFETAGASILLEFIPPELLIPVVGAPLL 672  
 Qy 935 S---DDKINEIEGPDSSRDSSENGRGD-----TTSNT---HDVR----- 966  
 Db 673 ESYDINKNKKIITDNLTKRNEKWSMDYGLIVAQWLSVTNTQVYTIKBMGYKALNYQAQ 732  
 Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHQYITQIEN--NG-----1003  
 Db 733 ALBHIKRYNIYSEKESNINIDFNDINSKLNQINGINQAINNFINFGCSVYLMKMI 792  
 Qy 1004 -----IIRGOEBSAGNSVNYKQPNKRSNFSSENDHKHKNIOEYNSR-----DTRKVRBEII 1053  
 Db 793 PLAVEKLLDFNTLKNLLNIDENKLYLIGSABEYKSVKNYKLTIMPFDLSIYTNDTI 852  
 Qy 1054 KLSQKNCNN 1063  
 Db 853 LIEMFNKYS 862

## RESULT 22

US-11-077-550-88  
 ; Sequence 88, Application US/11077550  
 ; Publication No. US20050244435A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shone, Clifford Charles  
 ; APPLICANT: Quinn, Conrad Padraig  
 ; APPLICANT: Foster, Keith Alan  
 ; APPLICANT: Chaddock, John  
 ; APPLICANT: Marks, Philip  
 ; APPLICANT: Sutton, J. Mark  
 ; APPLICANT: Stancome, Patrick  
 ; APPLICANT: Wayne, Jonathan  
 ; TITLE OF INVENTION: Recombinant Toxin Fragments  
 ; FILE REFERENCE: 1581.0130004  
 ; CURRENT APPLICATION NUMBER: US/11/077,550  
 ; CURRENT FILING DATE: 2005-03-11  
 ; PRIOR APPLICATION NUMBER: 10/241,596  
 ; PRIOR FILING DATE: 2002-09-12  
 ; PRIOR APPLICATION NUMBER: 09/255,829



Qy 469 TGTFTNKPCTGCPERRQTLCLGRYTLHHRGHEEDYKEHLGASIVYEAQLLYKYKKEKDE 528  
Db 216 ---NK-----GASIFNRKGYFSDPALILMHVHGL-YGKVVDD- 253  
Qy 529 NALCSIIIONS---YADLADIIGSDII-----KDYKKGME-----ENL 564  
Db 254 ---LPVNEKKFPMQSTDAIQAEELVTFGQDPSIITPSTDKSIYDKVLQNFPGIVDRL 310  
Qy 565 NKVNDKCKRNEESLKIFREKWD-----ENKENVWMSAVLKNKETCKDYKFKQKIPQF 619  
Db 311 NKVLVCTSDPNINININIKKFKDKYFVDESEGKYSI-----DVESFDKLYKS 358  
Qy 620 LRW-FKEW--GDDPCEKREKIYFESFKVECKKCCDENTCKNKCSEYKWKWIDLKSEY 676  
Db 359 LMFGETETNAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYTBEGFNISDKOM 412  
Qy 677 EKQVDKTKDKKKMYDNIDVENKKEANVYLKESKSE-----CKDVNFDKIFNESP 728  
Db 413 EKEYRGONKAINQAY---EISKEHLAVYKIQMCSAIEGRAPGICIDVNDLFFIADK 469  
Qy 729 NEYEDMCKCKDEIKYLNKIKYKPKTKHDIDYDITFSDTFGDTGPISINANINEQSGKDT 788  
Db 470 NSFDLCKSKNERIEYNTQSNYIENDFPINELILDIDL-----ISKIELPSENT 517  
Qy 789 NTGNSSETSDSPVSHESPDAAINVEKLSGDSESS-----SETRGILDINDPSVTNNVNEV 842  
Db 518 ESLTDFNVDPVVEKQPA-----IKKIFTDENTIFQVLYSQTFP-LDIRDISLTSSPDDA 571  
Qy 843 HDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSG-----NQSDN-----QENSS 893  
Db 572 LLFSNKVYSPFSDYI-----KTAN-----KVVEAGLFGWVKQIVNDFVIEANKS 617  
Qy 894 HSSDNGS-----SLTIGQVPSDNTQNTYDSON-----PHRDTNALASLP 934  
Db 618 NTMDKIADISLIVPYIGLALNVGNETAKGNFENAFETAGASILLEFIPPELLIPVVGAFLL 677  
Qy 935 S---DDKINIEGFDSDSRDSENGRD-----TTSNT---HDVR-----966  
Db 678 ESYIDNKKIKITIDNALTKENEKWSMDYGLIVAOMLSTVNTQFYTIKEGMYKALNQAQ 737  
Qy 967 -----RNVISERVNSHDFIRNGMAN--NNAHHOYITQIEN--NG-----1003  
Db 738 ALEBIIKYRYNIYSEKSNINIDFNDSKLNENGINQAIDNINNFINGCSVSLMKQMI 797  
Qy 1004 -----IIRGOEESAGNSVNYKPKRSPNSSENHKKNIQOYNSR-----DTKEVREII 1053  
Db 798 PLAVEKLDFTNLUKNLLNIDENKLYLGSAEYKSKVKNYKLTIMPFDLSIYTNDDTI 857  
Qy 1054 KLSKQNKCNN 1063  
Db 858 LIEMFNKYS 867

## RESULT 24

US-11-077-550-98  
; Sequence 98, Application US/11077550  
; Publication No. US2005024435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12

; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625596.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 98  
; TYPE: PRT  
; LENGTH: 867  
; ORGANISM: Clostridium botulinum  
US-11-077-550-98

Query Match 2.1%, Score 133.5; DB 7; Length 867;  
Best Local Similarity 18.6%; Pred. No. 0.098;  
Matches 180; Conservative 137; Mismatches 346; Indels 307; Gaps 45;  
Qy 252 TDRVKGIVNTKFSYDYKERNVEKLNINIKK-----EWEKKNANLNMHVMVNHKGNISKECA 307  
Db 47 TDRI-WIIPERYTFGYKPEDFNKSSGIFNRDVCYDYLN-----TNDKKNI-----94  
Qy 308 IIPAEEPQINLWIKENWENFLMEKRLFLNIDKCVENKKEACFGGC-----RLPCSS 361  
Db 95 -----FLOTMIKLFNRIKSKPLGEKLEMIINGIPIYLGDRRVLE 135  
Qy 362 YTSFMKSKTKQMEVLTNLKQKNSGVNDKNNFLNDLPKNNKNDLDDFFKNEKEYDDLDLDC 421  
Db 136 FN-----TWI-----ASVTNKLINPGEVERKKGI-----FAN-----164  
Qy 422 RYTATIKSFLNGPAKND---VDIASQINVNDLRGFG-----C-NYKS---NNEKSWNC 468  
Db 165 -----LIIFGPGVFLNENETIDIGIQNHFPASREGFGGIMQMFKFCEYVSFVNNVQE---215  
Qy 469 TGTFTNKPCTGCPERRQTLCLGRYTLHHRGHEEDYKEHLGASIVYEAQLLYKYKKEKDE 528  
Db 216 ---NK-----GASIFNRKGYFSDPALILMHVHGL-YGKVVDD- 253  
Qy 529 NALCSIIIONS---YADLADIIGSDII-----KDYKKGME-----ENL 564  
Db 254 ---LPVNEKKFPMQSTDAIQAEELVTFGQDPSIITPSTDKSIYDKVLQNFPGIVDRL 310  
Qy 565 NKVNDKCKRNEESLKIFREKWD-----ENKENVWMSAVLKNKETCKDYKFKQKIPQF 619  
Db 311 NKVLVCTSDPNININININIKKFKDKYFVDESEGKYSI-----DVESFDKLYKS 358  
Qy 620 LRW-FKEW--GDDPCEKREKIYFESFKVECKKCCDENTCKNKCSEYKWKWIDLKSEY 676  
Db 359 LMFGETETNAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYTBEGFNISDKOM 412  
Qy 677 EKQVDKTKDKKKMYDNIDVENKKEANVYLKESKSE-----CKDVNFDKIFNESP 728  
Db 413 EKEYRGONKAINQAY---EISKEHLAVYKIQMCSAIEGRAPGICIDVNDLFFIADK 469  
Qy 729 NEYEDMCKCKDEIKYLNKIKYKPKTKHDIDYDITFSDTFGDTGPISINANINEQSGKDT 788  
Db 470 NSFDLCKSKNERIEYNTQSNYIENDFPINELILDIDL-----ISKIELPSENT 517  
Qy 789 NTGNSSETSDSPVSHESPDAAINVEKLSGDSESS-----SETRGILDINDPSVTNNVNEV 842  
Db 518 ESLTDFNVDPVVEKQPA-----IKKIFTDENTIFQVLYSQTFP-LDIRDISLTSSPDDA 571  
Qy 843 HDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSG-----NQSDN-----QENSS 893  
Db 572 LLFSNKVYSPFSDYI-----KTAN-----KVVEAGLFGWVKQIVNDFVIEANKS 617  
Qy 894 HSSDNGS-----SLTIGQVPSDNTQNTYDSON-----PHRDTNALASLP 934  
Db 618 NTMDKIADISLIVPYIGLALNVGNETAKGNFENAFETAGASILLEFIPPELLIPVVGAFLL 677

```

Qy 935 S-----DDKINEIEGDSRDSSENGRGD-----TTSNT--HDVR----- 966
Db 678 ESYIDNKNKIITIDNALTKRNEKSDMYGLIWAQMLSTVNTQFYTIKEGMYKALNYQAQ 737
Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHGYITQIEN--NG----- 1003
Db 738 ALEBIIKYRNIYSEKSNINIDFNDSKLNBNQAINNFINNGSVSYLMKMKMI 797
Qy 1004 -----IIRGOESAGNSVNYKONPKRSNFSSENHDKNIQEVNSR-----DTKRVRBEII 1053
Db 798 PLAVEKLLDFDNTLKNLLNYIDENKLYLIGSAEYKSKVNYKLYKTIMPFDLSIYNTDTI 857
Qy 1054 KLSQKQKCNN 1063
Db 858 LIEMFNKYN 867

RESULT 25
US-11-077-550-92
; Sequence 92, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT FILING DATE: 2005-03-11
; PRIOR FILING DATE: 2005-03-11
; PRIOR FILING DATE: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR FILING DATE: 1999-02-23
; PRIOR FILING DATE: 1997-08-22
; PRIOR FILING DATE: 1996-12-27
; PRIOR FILING DATE: 1996-12-13
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
; LENGTH: 870
; TYPE: PRF
; ORGANISM: Clostridium botulinum
US-11-077-550-92

```

```

Query Match      2.1%; Score 133.5; DB 7; Length 870;
Best Local Similarity 18.6%; Pred. No. 0.098;
Matches 180; Conservative 137; Mismatches 346; Indels 307; Gaps 45;

Qy 252 TDRVKGINTKFSDYKKNVEKLNKKI-----EWEKKNKANLWNHNVHKGNIKECA 307
Db 50 TDRI-WIIPERYTGYKPEDFNKSGGFNRDVCYIDPDYLN-----TNDKKNI----- 97
Qy 308 IIPAEFPQINLWIKENWENFNLMEKKRLFLNLIKDKCVENKKYACFGGC-----RLPCSS 361
Db 98 -----FLQTMKLFNRIKSPGLCKLEMLINGIPYLGDERVPLEE 138
Qy 362 YTSFVKKSOTOMEVLTNLYKKNGSVKQNNFNLDFKKNQKNDLDPFKNEKEYDDLCDC 421
Db 139 FN-----TNI-----ASVTNKLISNPGVEVKKI---FAN----- 167
Qy 422 RYTATIISKPLNGPAKND---VDIASQINVDNLRGFG-----C-NYKS--NNEKSWNC 468

```

```

Db 168 -----LIIFGPGVPLNENETIDIGIONHFASREGFGGIMQMKFCPEYVSFVNNVOE--- 218
Qy 469 TGTFTNKPFGTCBPPRRQTLCLGRTYLLHRGHEEDYKEHLGASIYEAQLLLKTKYKEKDE 528
Db 219 -----NK-----GASIFNRGYSFSDPALILMHLELHVLHGL--GKIKVDD-- 256
Qy 529 NALCSIIQNS---YADLADIIGSDII-----KDYTKKKME-----ENL 564
Db 257 ---LPIVNPKEKPFMQSTDAIQABELYTFGGQDPSIITPSTDKSIYDKVLQNFGRVIDRL 313
Qy 565 NKYNKDKKNEESLKIIFREKWD-----ENKENVMKMSAVLKNKCKDYDFQKIPQF 619
Db 314 NKVLVCISDPNININIKNFKDKYKFVDESEKYSI-----DVESFDRUKYS 361
Qy 620 LRW-FKEW--GDDPCFKRKEIYSPESFKVCKKDCDCENTCKNKCKSEYKWKIDLKXSEY 676
Db 362 LMFGFTETNIAENYKIKTRASYFSDSLPVVKIK-----NLLDNEIYTIIEGFINISDKDM 415
Qy 677 EKQVCKYTKDKNKKMYDNIDEVKNKEANYLKEKSKE-----CKVNFDDKIFNESP 728
Db 416 EKEYRGONKAINKQAY---EISKEHLAVYKIQMCSAIEGRAPGICIDVDNEDLFFIADK 472
Qy 729 NEYEDMCKKCKDEIKYLNEIKYPKTKHDIVIDTFSDFGDTGPISINANINEQSGKDT 788
Db 473 NSFSDDLKSKNERIEYNTQSNYIENDPPINELIILDTDL-----ISKIELPSENT 520
Qy 789 NTGNSETSDSPVSHPEPESDAAINVEKLSGDESS-----SETRGIILINDPSTNNVNEV 842
Db 521 ESUTDFNVDVPVVEKQPA-----IKKIFTDENTIFQVLYSQTFP-LDIRDISLTSSPDDA 574
Qy 843 HDASNTQGSVNSWTSITNGHSESSLARTNNAQDIKIGRSG-----NEQSN-----QENSS 893
Db 575 LLFSNKVYSPFSMDYI-----KTAN-----KVVEAGLFAGWVKQIVNDFVIEANKS 620
Qy 894 HSSDMSG-----SLTIGQVPSEDTQNTYDSQN-----PHRDTNALASLP 934
Db 621 NTMDKIADISLIVPYIGLNLVGNETAKGNFENAFIAGASILELFIPELLIPVUGAFLL 680
Qy 935 S-----DDKINEIEGDSRDSSENGRGD-----TTSNT--HDVR----- 966
Db 681 ESYIDNKNKIITIDNALTKRNEKSDMYGLIWAQMLSTVNTQFYTIKEGMYKALNYQAQ 740
Qy 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHGYITQIEN--NG----- 1003
Db 741 ALEBIIKYRNIYSEKSNINIDFNDSKLNBNQAINNFINNGSVSYLMKMKMI 800
Qy 1004 -----IIRGOESAGNSVNYKONPKRSNFSSENHDKNIQEVNSR-----DTKRVRBEII 1053
Db 801 PLAVEKLLDFDNTLKNLLNYIDENKLYLIGSAEYKSKVNYKLYKTIMPFDLSIYNTDTI 860
Qy 1054 KLSQKQKCNN 1063
Db 861 LIEMFNKYN 870

RESULT 26
US-11-077-550-84
; Sequence 84, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT FILING DATE: 2005-03-11
; PRIOR FILING DATE: 2005-03-11
; PRIOR FILING DATE: 10/241,596

```





Db 622 NTMDKIADISLIVPIGLANVGNETAKGNFENAFIAGASILLPELPIPVVGAFLL 681  
 QY 935 S---DDKINEIEGDSRDSNGRGD-----TTSNT--HDVR----- 966  
 Db 682 ESYIDNKNKIITIDNALTKRNEKWSMDYGLIVAQWLSTVNTQYTIKEGMYKALANYQAQ 741  
 QY 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHQYITQIEN--NG----- 1003  
 Db 742 ALBIIKYRYNIYSEKEKSNINIDFNDINSKLNIGINQAIDNINNFINGCSVSLMKMI 801  
 QY 1004 -----IIRGOESAGNSVNYKDPKNSFSENDHKNIQIENSR-----DTRKRVREII 1053  
 Db 802 PLAVEKLLDFONTLKNLLNYIDENKLYLIGSAEYKSKVKNYKLTIMPFDLSIYTNDTI 861  
 QY 1054 KLSKQNKCN 1063  
 Db 862 LIEMFNKYS 871

## RESULT 29

US-11-077-550-171  
 ; Sequence 171, Application US/11077550  
 ; Publication No. US2005024435A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shone, Clifford Charles  
 ; APPLICANT: Quinn, Conrad Padraig  
 ; APPLICANT: Foster, Keith Alan  
 ; APPLICANT: Chaddock, John  
 ; APPLICANT: Marks, Philip  
 ; APPLICANT: Sutton, J. Mark  
 ; APPLICANT: Stancombe, Patrick  
 ; APPLICANT: Wayne, Jonathan  
 ; TITLE OF INVENTION: Recombinant Toxin Fragments  
 ; FILE REFERENCE: 1581.0130004  
 ; CURRENT APPLICATION NUMBER: US/11/077,550  
 ; PRIOR FILING DATE: 2005-03-11  
 ; PRIOR APPLICATION NUMBER: 10/241,596  
 ; PRIOR FILING DATE: 2002-09-12  
 ; PRIOR APPLICATION NUMBER: 09/255,829  
 ; PRIOR FILING DATE: 1999-02-23  
 ; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
 ; PRIOR FILING DATE: 1997-08-22  
 ; PRIOR APPLICATION NUMBER: 08/782,893  
 ; PRIOR FILING DATE: 1996-12-27  
 ; PRIOR APPLICATION NUMBER: GB9625996.5  
 ; PRIOR FILING DATE: 1996-12-13  
 ; PRIOR APPLICATION NUMBER: GB9617671.4  
 ; PRIOR FILING DATE: 1996-08-23  
 ; NUMBER OF SEQ ID NOS: 179  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 171  
 ; LENGTH: 862  
 ; TYPE: PRP  
 ; ORGANISM: Clostridium botulinum  
 US-11-077-550-171

Query Match: 2.0%; Score 132.5; DB 7; Length 862;  
 Best Local Similarity 18.5%; Pred. No. 0.11;  
 Matches 180; Conservative 137; Mismatches 346; Indels 309; Gaps 45;  
 QY 252 TDRVKGINTKFSYKKEKVKLNLIK-----EWEKGNKLANVHNVHKGNIKECA 307  
 Db 40 TDRI-WIPIRYTGYPKDPKNSGGIFNRDVCEYDPDYLN-----TNDKKN----- 87  
 QY 308 IIPAEPEQINLWIKWENEFNMEKRLFLNIDKDCVENKKYBACFGC-----RLPCSS 361  
 Db 88 -----FLOTWIKLFLNRKSKPLGEKLEMLINGIPYLGDRVPLER 128  
 QY 362 YTSFPMKSKTQMEVLTNLYKKNKSGVDKNNFLNDLPKKNQNDLDDPFKNEKEYDDLDC 421  
 Db 129 FN-----TNI-----ASVTNKLISNPGVEVERKGI---FAN----- 157

QY 422 RYTATIKSLPNGPAKND---VDIASQINVNDLRGFG-----C-NYKS--NNEKSWNC 468  
 Db 158 -----LIIIGPGPVLNENETIDIGIQNHFPASREGGIMQMFKCPYVSVFNVQOE--- 208  
 QY 469 TGTFTNKPOTCBPPRRQTLCLGRTYLLHRRGHEDYKEHLLGASIIYEAQLLKYYKKEKDE 528  
 Db 209 -----NK-----GASIFNRRGYFSDPALILMHILHVLHGL-YGIKYDD- 246  
 QY 529 NALCSIIQNS---YADLADIIGSDII-----KDYCKKKE-----ENL 564  
 Db 247 ---LPIVPNEKKFMGSTDAIQABELYFGQDPSIITPSTDKSIYDKVQNFGIVDRL 303  
 QY 565 NKYNKDKKRNESLKIIFREKWD-----ENKENVMKMSAVLKNKTKCHDYDFKQKIPQF 619  
 Db 304 NKVLVCISDPNININIKKFKDKYKFEVDESEKYSI-----DVESFDKLYKS 351  
 QY 620 LRW-FKEW--GDDPCEKREKIIYFBSFKVECKKQCDENTCKMKCSEYKWKIDLKSEY 676  
 Db 352 LMFGFTETNAENYKIKTRASYFSDSLPPVKIK-----NLLDNEIYITIEGFIISDKDM 405  
 QY 677 EKQVDKYTKDKNKKMYNDIDVENKKEANVYLKEKSE-----CKDVNFDKIFNE 726  
 Db 406 EKYRGONKAINKOAY---EBSIKEHLAVYKIQMCSAIEGRCSAPGICIDVDNEDLFIA 462  
 QY 727 SPNEYEDMKCKDBEIKYLNIEIKYKTKHDIYDITFSDTFDGDGTPISINANINEQQSGKD 786  
 Db 463 DKNSPDDLSKNERIEVNTQSYNIENDFPINELLIDTDL-----ISKIELPSE 510  
 QY 787 TSNTGNSETSDSPVSHPEPSDAAINVEKLSGDSS-----SETRGILDINDPSVTNNVN 840  
 Db 511 NTESLTDNFNVDPVYEKQPA-----IKKIPTDENTIPQYLYSQTFF-LDIRDISLTSFSD 564  
 QY 841 EVHDASTQGSVNTSDITNGHSESSLNRTNAQDIKGRSG-----NEQSDN-----QEN 891  
 Db 565 DALLFSNKVYSPFMDYI-----KTAN-----KVEAGLFGWVKQIVNDFVIEAN 610  
 QY 892 SSHSSDNGS-----SLATIGVPSRDNTQNTYDSON-----PHRDTPNALAS 932  
 Db 611 KSNYMDKIADISLIVPIGLANVGNETAKGNFENAFIAGASILLPELPIPVVGAF 670  
 QY 933 LPS---DDKINEIEGDSRDSNGRGD-----TTSNT--HDVR----- 966  
 Db 671 LLESYIDNKNKIITIDNALTKRNEKWSMDYGLIVAQWLSTVNTQYTIKEGMYKALNYQ 730  
 QY 967 -----RTNIVSERVNSHDFIRNGMAN--NNAHQYITQIEN--NG----- 1003  
 Db 731 AQALEBIIKYRYNIYSEKEKSNINIDFNDINSKLNIGINQAIDNINNFINGCSVSLMKK 790  
 QY 1004 -----IIRGOESAGNSVNYKDPKNSFSENDHKNIQIENSR-----DTRKRVRE 1051  
 Db 791 MIPLAVEKLLDFONTLKNLLNYIDENKLYLIGSAEYKSKVKNYKLTIMPFDLSIYTND 850  
 QY 1052 IIKLSKQNKCN 1063  
 Db 851 LIEMFNKYS 862

RESULT 30  
 US-11-077-550-167  
 ; Sequence 167, Application US/11077550  
 ; Publication No. US2005024435A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shone, Clifford Charles  
 ; APPLICANT: Quinn, Conrad Padraig  
 ; APPLICANT: Foster, Keith Alan  
 ; APPLICANT: Chaddock, John  
 ; APPLICANT: Marks, Philip  
 ; APPLICANT: Sutton, J. Mark  
 ; APPLICANT: Stancombe, Patrick  
 ; APPLICANT: Wayne, Jonathan  
 ; TITLE OF INVENTION: Recombinant Toxin Fragments  
 ; FILE REFERENCE: 1581.0130004  
 ; CURRENT APPLICATION NUMBER: US/11/077,550

```
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625596.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 167
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; US-11-077-550-167

Query Match          2.0%; Score 130.5; DB 7; Length 873;
Best Local Similarity 19.2%; Pred. No. 0.15;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

QY 388 DKWFL---NDLFKKNKNDLDDFFKNE-----KEYDDLCL-----DC 421
Db 85 EKONYLKVGVTLFRIYSTDLGRMLLTSIVRGIPFMGGSTIDTELKVIDTNCINVIQPDG 144
QY 422 RYTATIISFLNPGAKNDVDIASQINVDNLGRGFCNVKSNKSNWCT-----GTFNKK 475
Db 145 SYRSEELNVLIGSADIQF-----ECKSGHEVLNTRNGYSGTQVIRFSPDFTFG 197
QY 476 FPGTCEPRRQTLCLGR-----TYLLHRGHEEDYKEHLGLGASI-----513
Db 198 FEESLEVDWTPLLGAGKFPATPAVTLAHLQHLHAG-----RLYGIAINENRVFKYNTNA 251
QY 514 -YEAQLLYKYKE-----KDNALCSIIQNSYADLADII--KGSII-----552
Db 252 YEMSGLEVFSELRTPFGGHDAKFPIDSLQENEFRLYYNKFKDIASTLNKAISIVGTAS 311
QY 553 ----KDYGKK---MEENLNKVNKKRNEESLKIPREKWDENKENVKVMKMSAVLNKKE 605
Db 312 LQYKVNFKKYLISEBTSKFSVDKLFKLYKMLTEIYEDNFVKFPKVL-----NRK 366
QY 606 TKRDYD----KFQKIPQ-----FLRFKEWGGDDFCERKEKIYSPE 642
Db 367 TYLNFDKAVFKINIVPKVNTIYDGFNLRLNLAANFNGQNTENNNTFKLNFTGLPE 426
QY 643 SFKVEC-----KKKDCDENTCK---NKCSEYKKWIDLKKSBEYKQVDKYTKDKKNQMY 692
Db 427 FYKLLCVRGIIITSKTSLDRGYNKALMDLCIKVNNW--DLFFSPSE---DNFTDLNKGEE 482
QY 693 DNID-EVKNEANVYLKEKCEKDVNFDKIFNES--PNEYEDMKCKDCBIKYLNEIKYP 750
Db 483 ITSNTNIEAENISLDLQOYLYLTFFNDEPENISNIENLSSDIIIGOLEMP--NIERFP 540
QY 751 KTKHDIYDIDTFSDTGDGTPPISINANINQQSGKTSNTGNSSETSDSPVSHPEPSDAAI 810
Db 541 NGKK--YELDKYT-----MPLYLRAQFEHGKSKRIALTNS-----VNEALL 579
QY 811 NVEKLSGDESSSETRGTLINDPSVTNNVNEVDASNTQGSVN--TSDITNGHSESSLNR 869
Db 580 NPSRVYTFSSDYVK-----KVNKATEAAMFLGWVBQLVYDFDTESE--VST 625
QY 870 TTNAQDIKIGRSGNEQSDNGSSHSNDSGSLTIG-----OVPSEDNTQNTY 917
Db 626 TDKIADITIIPIYGPAINLGNMLYKDDDFVGAIFSGAVILLFIEIPIAIP-----VLGTF 681
QY 918 DSNQPHRDTNALASLPSSDKINEIEGDFSRRSSENGRGDTTNTNTHDVRRTNIVSERRVN 977
Db 682 ----ALVSVIA--NKVLTVTQTDIDNALSKRNEKWD---EVYKYIVTNWLA--KVN 724
```

```
QY 978 SH-DFIRNGMANNNAHQVITOENNGIIRGOESAGNSVNYKDNPKRNSFFSENHKKN 1036
Db 725 TQIDLIRKMKKE-----ALEN-----QAEATKAIINYQV--QYTEEEKNNINFN 767
QY 1037 IQEYNSRDTKRVREELIKLSK-QNKNNEYSMEYCTYSBERNSSPGPCSEERKKLCCOI 1095
Db 768 IDLSKLNESINKAMININKFLNQCVSYLM-----NSMIPYGVKRLIEDFDASL 817
QY 1096 SDYCLKY 1102
Db 818 KDALLKY 824

RESULT 31
US-11-074-176-152
; Sequence 152, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Kiaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; PRIOR FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 152
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
; US-11-074-176-152

Query Match          2.0%; Score 130; DB 7; Length 438;
Best Local Similarity 19.5%; Pred. No. 0.072;
Matches 90; Conservative 72; Mismatches 157; Indels 142; Gaps 19;

QY 560 MEENLNKVNKKRNEESLKIPREKWDENKENVKVMKMSAVLNKKTCKDYDKFQKIPQF 619
Db 19 MNSTIVHADKGTSHSISKVVSKTKND--KNVPESEQETSSNNIDQSQRKEK----72
QY 620 LRWFKEWGGDDFCERKEKIYSPEPKVECKKDCDENTCK---NKCSEYKKWIDLKXSEY 676
Db 73 -----EQAIPEQDQSQNTNNQDNDASEDEDEDEVSVEY 110
QY 677 EKQVDKYTKDKKNQMYNDIDEVKNKEANYVL-----KEKSKCKD-----V 717
Db 111 ENNVKDFRVKQVQVKDLAEKNNQHELMYIGRPTCYCQRFSPDLKDFNEIVKGKLLYF 170
QY 718 NFDKIKFNESPEYE-----DMCKCKDEIKYLE-----IKYPTKHDIYDIDTFSDTGF 767
Db 171 NIDDE---EGAHDYAFKVIPTPTPTFMNGKLISAWIGCEKGTQGLHDF--LYSDT--224
QY 768 DGTPTISINAN-INBQQSGKTSNTGNSSETSDSPVSHPEPSDAAINVEKLSGDESSSETRG 826
Db 225 -----ANKLVEQVVIKQNSNDTATQADNDVVASSEKTPTEVTVEENQAQSN-----271
QY 827 ILDINDPSVTNNV-NEV-HDASNTQGSVNTSDITNGHSE-----SSLNRTT-----871
Db 272 ----NDVAITNFAENSVPENAKNVASSTADLTQVATGDQDDVAPKAETKNTKTVKPKIHK 327
QY 872 -----NAQDIKIGRSGNEQSDNGSSHSNDSGSLTIGQVSEEDNTQNTYDS 919
Db 328 IVANKYKQAKLHKTNIIPMSAKKREDYKEN-----NOYDT 364
QY 920 QNPHRDTNALASLPSSDKINEIEGDFSRRSSENGRGDTTTS 960
Db 365 VKVHGTSFPAI-----KQKQARI---TMLKELENDTSDTIS 397
```

## RESULT 32

US-11-077-550-2  
; Sequence 2, Application US/11077550  
; Publication No. US2005024435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR FILING DATE: 2005-03-11  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 871  
; TYPE: PRT  
; ORGANISM: Clostridium botulinum  
US-11-077-550-2

Query Match 2.0%; Score 129.5; DB 7; Length 871;  
Best Local Similarity 19.2%; Pred. No. 0.17;  
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

Qy	388	DKNFL---NDLFKNNKNDLDDFFKNE-----KEYDDL-----DC	421
Db	83	EKNYLKGVTKLFRIYSTDGRMLTSIVRGIPFWGSGTIDELKVIDTNCINVIOPDG	142
Qy	422	RYTATIKSFLNGPAKNDVDIASQINVDNLRGFCNYSNNKSNWCT-----GTFTNK	475
Db	143	SYRSEELNLVIGPSADIQF-----ECKSFGHEVLNLRNGYGTQIRPSPDFTFG	195
Qy	476	PGTCBPFRQTLCLGR-----TYLLRHGHEEDYKEHLGASIT-----	513
Db	196	FEESLEVDTPNLLGAGKFAFDPAVTLAHLHAGH-----RLYGIAINPRVFKVTNA	249
Qy	514	-YEAQLKYYKKE-----KDNALCSIIQNSVADLADII-KGSDII-----	552
Db	250	YEMSGLEVSPEELRTFGHDAKPFIDSLQENEFRLYYNPKFOIATSLNKAISVGTAS	309
Qy	553	----KDYVGK---MEENLNKVNKKRNEBSLKIPEKWDENKNNVKNVSAVLNKE	605
Db	310	LOYNMNVPEKYLISEDSGKFSVDKLFKLYKMLTEIYTFEDNFVFKVL-----NRK	364
Qy	606	TKDQYD----KFQKIPQ-----FLRNPKEWGDDPCEKKEKIYSPE	642
Db	365	TYLNFDAVAFKINIVPKVNYTIYDGFNLRLNTLAANFNGQNTNMMNFTKLNFTGLFE	424
Qy	643	SPKVEC-----KKQCDENTCK---NKCSEYKKWIDLKSEYKQVDKTKDKNNKY	692
Db	425	FYKLLCVRGIIITSKIDKGNKALNDLCIKVNNW-DLFFSPSE---DNFTDLNKGEE	480
Qy	693	DNID-EVKNKEANVYLKESKECKDVNFDKIFNES-PNEYEDMKCKDEIKYLINEIKYP	750

Db	481	ITSNTNIEAAENISLDLIQYYLTFNFDNPEPENISLENLSSDIIGOLELMP--NIERRFP	538
Qy	751	KTKHDIYDITFSDTFGDTGTPISINANINEQOQSGKDTNTGNSETS DSPVSHPESDAAI	810
Db	539	NGKK--YELDKYT-----MPLYLRAQEFHKGSRIALTNS-----VNALL	577
Qy	811	NVEKLSGDESSSETRGILDINDPSVTNNVNEVHDASNTQGSVN--TSDITNGHSESSLNR	869
Db	578	NPSRVVTFSSDYVK-----KVNKATEAAMFLGWQLVYDFDTETSE--VST	623
Qy	870	TTNAQDIKIGRSGNEQSDNQSNSHSDNSGSLTIG-----QVPSEDNTONTY	917
Db	624	TDKIADITIIPIYGPAIINGNMLYKDDDFGALIFSGAVILLEPIPIAIP-----VLGTG	679
Qy	918	DSQNPREDTPNALASLPDDKINEIEGFSRSDSNGRGDTTNTDVRRTNIVSERRVN	977
Db	680	-----ALVSYIA-NKVLTVQIDNALSKRNEKW-----EVYKIYTNWLA--KVN	722
Qy	978	SH-DFIRNGMANNNAHHQYITQIENNGIIRGOESAGNSVNYKDNPKRSNFFSENDHKCN	1036
Db	723	TQIDLIRKMKKE-----ALEN-----QAEATKAIINYQV--QYTEEEKNNINFN	765
Qy	1037	IQEYNSRDTKRVREIEIKLSK-QNKNNEYSMEYCTYSDERNSSPGPCSRERKLCQCI	1095
Db	766	IDLSSKLNESINKAMININKFLNQCSVYLM-----NSMIPYGVKRLDFDASL	815
Qy	1096	SDYCLKY 1102	
Db	816	KDALLKY 822	

## RESULT 33

US-11-077-550-8  
; Sequence 8, Application US/11077550  
; Publication No. US2005024435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 871  
; TYPE: PRT  
; ORGANISM: Clostridium botulinum  
US-11-077-550-8

Query Match 2.0%; Score 129.5; DB 7; Length 871;  
Best Local Similarity 19.2%; Pred. No. 0.17;  
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

Qy 388 DKNFL---NDLFKNNKNDLDDFFKNE-----KEYDDL-----DC 421

```
Db      83 EKNYLKGVTKLPERIYSTDLGRMLTISVIRGIPFWGGSTIDTELKVIDTNCINVTQPDG 142
Qy      422 RYATATIKSFLNGPAKNDVDIASQINVDLRGFCGNYSNNKSNWCT-----GTFTNK 475
Db      143 SYRSEELNLVIIGPSADIIQF-----ECKSFGHEVLNLRNGYSGTQVIRFSPDFTFG 195
Qy      476 FPGTCPPRQTLCLGR-----TYLLHRGHEEDYKEHLLGASI-----513
Db      196 FBESLEVDTNPLGAGKFPATPAVTLAHLIHAH-----RLYGAINRNRVFKVNTNA 249
Qy      514 -YEAQLLKYYKE-----KDNALCSIIQNSYADLADI--KGSDII-----552
Db      250 YEMSGLEVSFEELRTFGGHDAKFDLSQENFRLYYNFKDIASTLNKAKSIGVGTAS 309
Qy      553 -----KDYGGK---MEENLNKVNKKRNEESLKIFREKWMNDENKENVKMSAVLNKE 605
Db      310 LQYMKNVFKEKYLSEDTSGKFSVDLKFDPKLYKMLTEIYTDNFVKFFKVL-----NRK 364
Qy      606 TKDQYD---KFQKIPQ-----FLRFKKGDDDFCEKRKEKISYFE 642
Db      365 TYLNFDAKVFKNIVPKVNYTIYDGFNLNRLNTLNAANFNQNTENNMFNFKLNFTGLFE 424
Qy      643 SFKVEC-----KKQDCDENTCK---NKCSEYKKWIDLKSEYKQVQDKYTKDKNKKMY 692
Db      425 FYKLLCVRGIIITSKYSKLDGYNKALNDLCIKNNW--DLFFSPSE---DNFTDNLKNGEE 480
Qy      693 DNID-EVKNKEANVYLKESKECKOVNFDKIPNES--PNEYEDMCKCKDEIKYLNKIKYP 750
Db      481 ITSDTNIEAAEENISLDLIQYYLTFNFDNENPENISLENLSSDIIGQLELMP--NIERFP 538
Qy      751 KTKHDIYDITFSDTGDGTPISINANINEQQSGKDTSTNGNSETSDSPVSHPEPESDAAI 810
Db      539 NGKK--YELDKYT-----MPLYRAQEFHGKSRIALTNS-----VNEALL 577
Qy      811 NVEKLSGDESSSTRGILDINDPSTNNVNVNVDASNTQGSVN--TSDITNGHSESSLNR 869
Db      578 NPSRVYTFSSDYVK-----KVNKATEAMFLGWVQLVYDFTDEISE--VST 623
Qy      870 TTNAQDIKIRGSGNEQSDNOSHSSDNGSLTIG-----QVPSEDNNTQNTY 917
Db      624 TDKIADITIIIPYGPALNIGNMLYKDDFVGALIFSGAVILLFPIEIAIP-----VLGTF 679
Qy      918 DSQNPHRDTPNALASLSDKINEIEGDFDSRDSSENGRGDTTNTHTDVRTNIVSERVNV 977
Db      680 -----ALVSYIA--NKVLTVQTDIDNALSKEKWD---EVYKYIVTNWLA--KVN 722
Qy      978 SH--DFIRNGMANNNAHHQYITQIENNGIIRGOESAGNSVNYKDNPKRSNFSSENHKKN 1036
Db      723 TQIDLRKKWKE-----ALEN-----QAEATKAIINYQN--QYTEEBKNNINFN 765
Qy      1037 IQBYSNDRTKRBEIIFKLSK-QMKCNNEYSMEYCTYSDBRNSSPGFCSEERERKGLCCQI 1095
Db      766 IDLLSKLINESINKAMININKFNQCSVSLM-----NSMIPYGVKRLDFDASL 815
Qy      1096 SDYCLKY 1102
Db      816 KDALLKY 822
```

## RESULT 34

```
US-11-077-550-26
; Sequence 26, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Radraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
```

```
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patentn version 3.1
; SEQ ID NO 26
; LENGTH: 871
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-26
```

```
Query Match      2.0%; Score 129.5; DB 7; Length 871;
Best Local Similarity 19.2%; Pred. No. 0.17;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;
```

```
Qy      388 DKNFEL---NDLFKKKNKNDLDDFFPKNE-----KEYDDLC-----DC 421
Db      83 EKNYLKGVTKLPERIYSTDLGRMLTISVIRGIPFWGGSTIDTELKVIDTNCINVTQPDG 142
Qy      422 RYATATIKSFLNGPAKNDVDIASQINVDLRGFCGNYSNNKSNWCT-----GTFTNK 475
Db      143 SYRSEELNLVIIGPSADIIQF-----ECKSFGHEVLNLRNGYSGTQVIRFSPDFTFG 195
Qy      476 FPGTCPPRQTLCLGR-----TYLLHRGHEEDYKEHLLGASI-----513
Db      196 FBESLEVDTNPLGAGKFPATPAVTLAHLIHAH-----RLYGAINRNRVFKVNTNA 249
Qy      514 -YEAQLLKYYKE-----KDNALCSIIQNSYADLADI--KGSDII-----552
Db      250 YEMSGLEVSFEELRTFGGHDAKFDLSQENFRLYYNFKDIASTLNKAKSIGVGTAS 309
Qy      553 -----KDYGGK---MEENLNKVNKKRNEESLKIFREKWMNDENKENVKMSAVLNKE 605
Db      310 LQYMKNVFKEKYLSEDTSGKFSVDLKFDPKLYKMLTEIYTDNFVKFFKVL-----NRK 364
Qy      606 TKDQYD---KFQKIPQ-----FLRFKKGDDDFCEKRKEKISYFE 642
Db      365 TYLNFDAKVFKNIVPKVNYTIYDGFNLNRLNTLNAANFNQNTENNMFNFKLNFTGLFE 424
Qy      643 SPKVEC-----KKQDCDENTCK---NKCSEYKKWIDLKSEYKQVQDKYTKDKNKKMY 692
Db      425 FYKLLCVRGIIITSKYSKLDGYNKALNDLCIKNNW--DLFFSPSE---DNFTDNLKNGEE 480
Qy      693 DNID-EVKNKEANVYLKESKECKOVNFDKIPNES--PNEYEDMCKCKDEIKYLNKIKYP 750
Db      481 ITSDTNIEAAEENISLDLIQYYLTFNFDNENPENISLENLSSDIIGQLELMP--NIERFP 538
Qy      751 KTKHDIYDITFSDTGDGTPISINANINEQQSGKDTSTNGNSETSDSPVSHPEPESDAAI 810
Db      539 NGKK--YELDKYT-----MPLYRAQEFHGKSRIALTNS-----VNEALL 577
Qy      811 NVEKLSGDESSSTRGILDINDPSTNNVNVNVDASNTQGSVN--TSDITNGHSESSLNR 869
Db      578 NPSRVYTFSSDYVK-----KVNKATEAMFLGWVQLVYDFTDEISE--VST 623
Qy      870 TTNAQDIKIRGSGNEQSDNOSHSSDNGSLTIG-----QVPSEDNNTQNTY 917
Db      624 TDKIADITIIIPYGPALNIGNMLYKDDFVGALIFSGAVILLFPIEIAIP-----VLGTF 679
Qy      918 DSQNPHRDTPNALASLSDKINEIEGDFDSRDSSENGRGDTTNTHTDVRTNIVSERVNV 977
```

Db 680 -----ALVSYIA-NKVLTVQTDNALSKRNEKWD-----EVYKYIVTNWLA--KVN 722  
Qy 978 SH-DFTIRNGMANNHQQYITQIENNGIIRGOESAGNSVNYKDNPKRSNFSSENDHKK 1036  
Db 723 TQIDLIRKKWKE-----ALEN-----QAEATKAIINYQYN--QYTEEEKNNINFN 765  
Qy 1037 IQEYNSRDTKRVREIILKSK-QNKCNNSEYCTYSDBRNSPSPGCSREERKCLCCQI 1095  
Db 766 IDLLSSKLNESINKAMININKFLNQC SVSLM-----NSMIPYGVKRLDPDASL 815  
Qy 1096 SDYCLKY 1102  
Db 816 KDALLKY 822

RESULT 35  
US-11-077-550-153  
; Sequence 153, Application US/11077550  
; Publication No. US2005024435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 153  
; LENGTH: 871  
; TYPE: PR1  
; ORGANISM: Clostridium botulinum  
US-11-077-550-153

Query Match 2.0%; Score 129.5; DB 7; Length 871;  
Best Local Similarity 19.2%; Pred. No. 0.17;  
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

Qy 388 DKNFL---NDLFKNNKNDLDDFFKNE-----KEYDDL---DC 421  
Db 83 EKONYLKGVTKLPERIYSTDLGRMLTISIVRGIPFWGGSTIDTELKVIDTNCINVIQPDG 142  
Qy 422 RYTATIISFLNGPAKNDVDIASQINVDLGRGCNYSKNEKSWNCT-----GTFN 475  
Db 143 SYRSEELNLVIIGSADIIOF-----ECKSFGHEVLNTRNGYGSTQVIRFSPDFTFG 195  
Qy 476 FPGTCEPRPOTLCLGR-----TYLLRHGHEBDYKEHLIGASI----- 513  
Db 196 FESLEVDVTPPLGAGKEATDPATLAHELHAGH-----RLYGIAINRNVFKVNTNA 249  
Qy 514 -YEAQLKLYKYE-----KDNALCSITONSADLADII-KGS011----- 552  
Db 250 YWMSGLEVSFEELRTGCGHDAKFDLSQENEFRLYYNFKDIATSLNKAISIVGTAS 309

Qy 553 ----KDYGGK---MBENLNKVNKDKRNEESLKIIFREKWMNDENKENVKVMGSAVLKNKE 605  
Db 310 LQYMKVPEKYLSSDTSKFSVDKLPDKLYKMLTEIYTEDNFVKVFKVL-----NRK 364  
Qy 606 TKDYD----KFKIPIQ-----FLWFKWGGDDFCBKRKEKYSFE 642  
Db 365 TYLNFDPKAVKINIVPKVNYTYIDGFNLRTNLAAFNQONTINNMFYTKLNKFTGLFE 424  
Qy 643 SPVEEC-----KKDCDENTCK---NKCSEYKKWIDLKKSEVEKQVDKYTKDKNKM 692  
Db 425 FYKLLCVRGIIITSKSLDKGYNKALNDLCIKVNW-DLFFSPSE---DNFTDNLKGE 480  
Qy 693 DNID-EVYKKBANVYLKSKSKECKVNFDDKIFNES-PNEYEDVCKCDBIKYLNKIKYP 750  
Db 481 ITSDTNIEAAENISLDLIQQYVLTFFNFONPENISLENLSSDIIGQLELMP-NIERFP 538  
Qy 751 KTXHDIYDIDTFSDTFGDGTPIISINANINEQSGKDTSTNGNSETSDSPVSHPEPESDAI 810  
Db 539 NGKK--YELDKYT-----MFWLRAQEFHKGSRIALTNS-----VNSALL 577  
Qy 811 NVEKLSGDESSSSTRGILDINDPSVTNNVNEVDASNTQGSVN-TSDITNGHSESSLNR 869  
Db 578 NPSRVYTFSSDVYK-----KVNKATEAMFLGWVEQLVYDFTDSE--VST 623  
Qy 870 TTNAQDIKIGRSGNEQSDNQENSHSDNSGSLTIG-----QVPSSEDNTONTY 917  
Db 624 TDKIADITIIPIYGPALNIGNMLYKDDFVGALIFSGAVILLFPIEIAIP---VLGTF 679  
Qy 918 DSQNPDRPTNALASLPDDKINEIGSPDSRSENGRGTNTNTHDVRNTNIVSERVN 977  
Db 680 -----ALVSYIA-NKVLTVQTDNALSKRNEKWD---EVYKYIVTNWLA--KVN 722  
Qy 978 SH-DFTIRNGMANNHQQYITQIENNGIIRGOESAGNSVNYKDNPKRSNFSSENDHKK 1036  
Db 723 TQIDLIRKKWKE-----ALEN-----QAEATKAIINYQYN--QYTEEEKNNINFN 765  
Qy 1037 IQEYNSRDTKRVREIILKSK-QNKCNNSEYCTYSDBRNSPSPGCSREERKCLCCQI 1095  
Db 766 IDLLSSKLNESINKAMININKFLNQC SVSLM-----NSMIPYGVKRLDPDASL 815  
Qy 1096 SDYCLKY 1102  
Db 816 KDALLKY 822

RESULT 36  
US-11-077-550-6  
; Sequence 6, Application US/11077550  
; Publication No. US2005024435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4

; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; TYPE: PRT  
; LENGTH: 873  
; ORGANISM: Clostridium botulinum  
US-11-077-550-6

Query Match  
Best Local Similarity 19.2%; Score 129.5; DB 7; Length 873;  
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

```
QY 388 DKNFL---NDLFKNNKNDLDDFFKNE-----KEYDDL---DC 421
Db 85 EKONYLKVGTGFERIYSTDLGRMLTISVIRGIPFGGSTDITELKVIDTNCINVIQPDG 144
QY 422 RYTATTIKSFLNGPAKNDVDIASQINVDLGRGFCNYKSNKSNWCT-----GTFTNK 475
Db 145 SYRSEELNLVIIGPSADIQF-----ECKSFGHEVLNLRNGYGSQVIRSPDFTFG 197
QY 476 FPGTCPPRRQTLCLGR-----TYLLHRGHEEDYKEHLLGASI-----513
Db 198 FEESLEVDTPNLLGAGKFAFDPAVTLAHLIHAH-----RLYGIAINPNRVKVTNA 251
QY 514 -YEAQLLKYYKE-----KDNALCSIIQNSYADLADI--KGSII-----552
Db 252 YEMSGLEVSFEELRTFGGHDKAFIDSLQENFRLYYNKFQDIASITLNKAKSIVGTAS 311
QY 553 ----KDYGGK---MEENLNKVNKDKRNEESLKIPEKWMNDENKENVKMSAVLNKE 605
Db 312 LQYMKNVFKEKYLSEDTSGKFSVDKLPDKLYKMLTEIYEDNFVKFFVL-----NRK 366
QY 606 TKDYPD---KFOKIPQ-----FLRFKMGDDFCRKRKEKISYE 642
Db 367 TYLNFDAKAVKINIVPKVNTIYDGNLNTLAANFNGQNTENNNTFKLNFTGLPE 426
QY 643 SFKVEC-----KKDCDENTCK---NKCSEYKWKIDLKSEYKQVDKYTKDKNKMY 692
Db 427 FYKLLCVRGIIITKTSKSLDKGYNKALDLCIKVNNW--DLFFSPSE---DNFTDLNKGEE 482
QY 693 DNID-EVKKNEANVYLKESKECKVDNFDDKIFNES--PNEYDMCKCKDEIKYLNKYP 750
Db 483 ITSDTNIEAAEENISLDLIQYYLTFNFDNEPENISNIENLSSDIIGGLELMP--NIERFP 540
QY 751 KTKHDIYDITFSDTFGDGTPIINANINQOQSGKOTSGTNSGSETSDSPVSHPEPSDAI 810
Db 541 NGKK--YELDKYT-----MFWYLAQSFEGHKSRIALTNS-----VNEALL 579
QY 811 NVEKLSDGESSSETRGILDINDPSVTNNVNEVHDASNTQSVSN--TSDITNGHSESLNR 869
Db 580 NPSRVYTFSSDYVK-----KVNKATEAAMFLGWVQLVYDFDTSB--VST 625
QY 870 TTNAQDIKIGRSGNEQSDNQSNSHSDNSGSLTIG-----QVPSDNTQNTY 917
Db 626 TDKIADITIIPIYGPAIINLGMWLYKDDFVGALIFGCAVILLBFIPIAIPT-----VLGTF 681
QY 918 DSQNPHRDTPNALASLPDDKINEIEGFSRDSRSENGRGDTTNTNHDVRRNTNIVSERRVN 977
Db 682 -----ALVSYIA-NKVLTVQTDNALSKEKWD---EVYKIYITNMLA--KVN 724
QY 978 SH-DPIRNGMANNNAHQYITOENNGIIRGOESAGSNVYKDNPKRSPFSENHDKN 1036
Db 725 TQIDLIRKKWKE-----ALEN-----QAEATKAIINYQN--QYTEERKNINFN 767
QY 1037 IQSYNSRDTKRVREIITKLK-OMKCNNEYSMEYCTYSDBRNSPGPCSREREKLCQOI 1095
Db 768 IDLSSKLNESINKAMININKFLNQCSVSLM-----NSMIPYGVKRLIEDFASL 817
QY 1096 SYCLKY 1102
Db 818 KDALLKY 824
```

## RESULT 37

US-11-077-550-149  
; Sequence 149, Application US/11077550  
; Publication No. US20050244435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 149  
; TYPE: PRT  
; LENGTH: 873  
; ORGANISM: Clostridium botulinum  
US-11-077-550-149

Query Match 2.0%; Score 129.5; DB 7; Length 873;  
Best Local Similarity 19.2%; Pred. No. 0.17;  
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

```
QY 388 DKNFL---NDLFKNNKNDLDDFFKNE-----KEYDDL---DC 421
Db 85 EKONYLKVGTGFERIYSTDLGRMLTISVIRGIPFGGSTDITELKVIDTNCINVIQPDG 144
QY 422 RYTATTIKSFLNGPAKNDVDIASQINVDLGRGFCNYKSNKSNWCT-----GTFTNK 475
Db 145 SYRSEELNLVIIGPSADIQF-----ECKSFGHEVLNLRNGYGSQVIRSPDFTFG 197
QY 476 FPGTCPPRRQTLCLGR-----TYLLHRGHEEDYKEHLLGASI-----513
Db 198 FEESLEVDTPNLLGAGKFAFDPAVTLAHLIHAH-----RLYGIAINPNRVKVTNA 251
QY 514 -YEAQLLKYYKE-----KDNALCSIIQNSYADLADI--KGSII-----552
Db 252 YEMSGLEVSFEELRTFGGHDKAFIDSLQENFRLYYNKFQDIASITLNKAKSIVGTAS 311
QY 553 ----KDYGGK---MEENLNKVNKDKRNEESLKIPEKWMNDENKENVKMSAVLNKE 605
Db 312 LQYMKNVFKEKYLSEDTSGKFSVDKLPDKLYKMLTEIYEDNFVKFFVL-----NRK 366
QY 606 TKDYPD---KFOKIPQ-----FLRFKMGDDFCRKRKEKISYE 642
Db 367 TYLNFDAKAVKINIVPKVNTIYDGNLNTLAANFNGQNTENNNTFKLNFTGLPE 426
QY 643 SFKVEC-----KKDCDENTCK---NKCSEYKWKIDLKSEYKQVDKYTKDKNKMY 692
Db 427 FYKLLCVRGIIITKTSKSLDKGYNKALDLCIKVNNW--DLFFSPSE---DNFTDLNKGEE 482
QY 693 DNID-EVKKNEANVYLKESKECKVDNFDDKIFNES--PNEYDMCKCKDEIKYLNKYP 750
Db 483 ITSDTNIEAAEENISLDLIQYYLTFNFDNEPENISNIENLSSDIIGGLELMP--NIERFP 540
```



```
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 159
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; US-11-077-550-159

Query Match
Best Local Similarity 2.0%; Score 129.5; DB 7; Length 879;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

Qy 388 DKNFL---NDLFKNKNDLDDFFKNE-----KEYDLC-----DC 421
Db 83 EKONYLKVTKLPERIYSTDLGRMLTISIVGIPFGSGTIDTDLKVIDTNCINVIQPDG 142
Qy 422 RYATIIKSLFNGPAKNDVDIASQINNDLGRGFCNYKSNKSNWCT-----GTPTNK 475
Db 143 SYRSEELNLVIIGPSADIIQF-----ECKSFGHEVLNLRNGYSGTQVIRFSPDFTFG 195
Qy 476 FPGTCPPRRQTLCLGR-----TYLLHRGHEEDYKEHLLGASI-----513
Db 196 FESLEVDVTPNLLGAGKFAFDPAVTLAHELHAGH-----RLYGTAINPNRVFKVNTNA 249
Qy 514 -YEAQLLKYYKE-----KDNALCSIIQNSYADLADII-KGSDII-----552
Db 250 YEMSGLEVSFEELRTFGGHDAKFIDSLQENEFPLYYNFKDIASLTKAKSIVGTAS 309
Qy 553 ----KDYGKK---MEENLNKVNKKRNEESLKI-FREKWDENKENVKMSAVLKKE 605
Db 310 LOYMKNVFEKYLSEDTSGKFSVDKLFKDKLYKMLTEIYTDNFVKFFKVL-----NRK 364
Qy 606 TCKDYD---KFQKIPO-----FLRWFKWGDDFCERKEKISFE 642
Db 365 TYLNFDAVFKINIVPKVNTIYDGFNLNLTNLAANFNGQNTNINNFTKLKNFTGLFE 424
Qy 643 SFKVEC-----KKDCDENTCK---NKCSEYKKWIDLKKSEYKQVDKYTKDKNKMV 692
Db 425 FYKLLCVRGIIITSKTSLDKYGNKALDLCIKVNMW-DLFFSPSE---DNFTNDLNKGE 480
Qy 693 DNTD-EVKNKEAVYLLKEKCKVDNFDKIFNES-PNEYEDMCKKCDIKYLNKYP 750
Db 481 ITSDTNTAEABENISLDLIQOYLYLTFFNDEPENISIENTLSSDIIGOLELMP--NIERFP 538
Qy 751 KTKHDYIDIDTDFDFTGDTGPISINANIEQSGKDTNTGNSSETSDSPVSHPESDAAI 810
Db 539 NGKK--YELDKYT-----MFHLRAQEFHGKSRIALTNS-----VNEALL 577
Qy 811 NVEKLSGDESSSETRGILDINDPSVTNNVNEVHDASNTQGSVSN-TSDITNGHSESLNR 869
Db 578 NPSRVYTFSSDYVK-----KVNKATEAMFLGWVQVVDFTDETSE--VST 623
Qy 870 TTHAQDTIKGRSGNEQSDNENSHSDNSGSLTIG-----QVPSEDNTQNTY 917
Db 624 TDKIADITIIPIYTGPAIINGNMLYKDDFVGALIFSGAVILLEFIPFIAIP-----VLGTF 679
Qy 918 DSQNPHRDTPNALASLPDDKINIEGFDSDSRDSRDSGRGDTTSTNTHDVRRTNIVSERVN 977
```

```
Db 680 -----ALVSVIA-NKVLTVQTIIDNALSKRNEKWD-----EVYKYIVTNMLA--KVN 722
Qy 978 SH-DFIRNGMANNNAHQVITQIENNGIIRGOBESAGSNVYKDNPKRSNFSSENHKKN 1036
Db 723 TQIDLLRKKE-----ALEN-----QAEATKALINYQN--QYTEEEKNNINFN 765
Qy 1037 IQEYNSRDYKRVREELIKLSK-ONKCNNEYSMEYCYTYSBERNSPPGSPCSREERKKLCCQI 1095
Db 766 IDDLSSKLNESINKAMININKFLNQCYSVLM-----NSMIPYGVKRLDFDASL 815
Qy 1096 SDYCLAY 1102
Db 816 KDALLKY 822

RESULT 40
US-11-077-550-4
; Sequence 4, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; US-11-077-550-4

Query Match
Best Local Similarity 2.0%; Score 129.5; DB 7; Length 894;
Matches 163; Conservative 123; Mismatches 322; Indels 239; Gaps 40;

Qy 388 DKNFL---NDLFKNKNDLDDFFKNE-----KEYDLC-----DC 421
Db 106 EKONYLKVTKLPERIYSTDLGRMLTISIVGIPFGSGTIDTDLKVIDTNCINVIQPDG 165
Qy 422 RYATIIKSLFNGPAKNDVDIASQINNDLGRGFCNYKSNKSNWCT-----GTPTNK 475
Db 166 SYRSEELNLVIIGPSADIIQF-----ECKSFGHEVLNLRNGYSGTQVIRFSPDFTFG 218
Qy 476 FPGTCPPRRQTLCLGR-----TYLLHRGHEEDYKEHLLGASI-----513
Db 219 FESLEVDVTPNLLGAGKFAFDPAVTLAHELHAGH-----RLYGTAINPNRVFKVNTNA 272
Qy 514 -YEAQLLKYYKE-----KDNALCSIIQNSYADLADII-KGSDII-----552
Db 273 YEMSGLEVSFEELRTFGGHDAKFIDSLQENEFPLYYNFKDIASLTKAKSIVGTAS 332
Qy 553 ----KDYGKK---MEENLNKVNKKRNEESLKI-FREKWDENKENVKMSAVLKKE 605
```

```

Db      333  LQYMKNVFKYLLSEDTSGKFSVDKLFKXKMLTEIYTEDNFVKFFKVL-----NRK 387
Qy      606  TCKDYD---KFQKIPQ-----FLRWFKEWGDDPCEKKEKIYSFE 642
Db      388  TYLNFDAVKAVKINIVPKVNYIYDGFNLRNTNLAANFNGQNTENNWNFTLKNKFTGLFE 447
Qy      643  SFKVEEC-----KKKCDENTCK---NKCSEYKKWIDLKSEYKQVDKYTKDKNKKMY 692
Db      448  FYKLLCVRGIIITSKTSKLDKGNKALNDLCIKVNNW-DLFFPSPSE---DNFTNDLNKGE 503
Qy      693  DNID-EVKNKEANYLKEKCEKDVNDDKIFNES-PNEYEDMCKCKDEIKYLNEIKYP 750
Db      504  ITSDTNTEAABENISLDLIQYYLTFNFDNBPENISIEENLSSDIIGOLELMP--NIEFP 561
Qy      751  KTKHDIYDIDTFDGTGTPISINANINEQQSGKDTSTNCGNSETSDSPVSHPEPESDAAI 810
Db      562  NGKK--YELDKYT-----MFHYLRAQBFHKGKSRALTNS-----VNEALL 600
Qy      811  NVEKLSGDESSSETRGILDINDPSTVNNVNEVHDASNTQGSVSN-TSDITNGHSESSINR 869
Db      601  NPSRVYTFSSDYVK-----KVKATEAAMFLGWVEQLVYDFTDETSE--VST 646
Qy      870  TTNAQDIKIGRSGNEQSDNQENSHSDNSGLTIG-----QVPSEDNTQNTY 917
Db      647  TDKIADITIIIPYIGPALNIGMMLYKODFVGALLIFSGAVILLEFPEIAP-----VLGTF 702
Qy      918  DSQNPHRDTPNALASLPDDKINIEGPDSSRDSRSGRGDTSTNTHDVRRTNIVSERRVN 977
Db      703  -----ALVSYIA-NKVLTVQTIIDNALSRRNEKWD---EVYKIVTNWLA--KVN 745
Qy      978  SH-DFIRNGMANNAHQYITQIENNGIIRQGESAGNSVNYKDNPKRSNFSSENDRHKN 1036
Db      746  TQIDLIRKKMKE-----ALEN-----QAEATKALINYQYN--QYTEEEKNNINFN 788
Qy      1037 IQEYNSRDTKRVREIILKSK-QNKCNNEYSMEYCTYSDERNSSPGPCSRERKKLCCQI 1095
Db      789  IDDLSSKLINESINKAMININKFLNQCSVSYLM-----NSMIPYGVKRLDFDASL 838
Qy      1096 SDYCLKY 1102
Db      839 KDALLKY 845

```

Search completed: November 21, 2005, 20:52:45  
Job time : 30 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2005, 20:17:17 ; Search time 193 Seconds  
(without alignments)  
2754.655 Million cell updates/sec

Title: US-10-677-980-2

Perfect score: 6481

Sequence: 1 MKGYFNLYFLPLFLYLNVI.....VQETNISDSYNNKMMY 1210

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_21.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*
- 9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6481	100.0	1210	6	ABG73547 P. falcip
2	6081	93.8	1143	5	Aau76759 Plasmodiu
3	1557	24.0	1435	2	AAR70232 P. falcip
4	1557	24.0	1435	2	Aaw22477 Silaia ac
5	1557	24.0	1435	3	Aay77900 P. falcip
6	1543.5	23.8	1604	2	Aar70105 TNF-R-EBA
7	1542.5	23.8	1786	2	AAR41043 CD4-EBA17
8	1537.5	23.7	1421	5	Aau76764 Plasmodiu
9	1200.5	18.5	616	6	ABR82499 Synthetic
10	1199.5	18.5	616	6	ABR82498 Synthetic
11	1197.5	18.5	616	5	AAM50533 Unidentif
12	1112	17.2	1086	5	Aau76760 Plasmodiu
13	1088	16.8	1568	6	ABP70152 Amino aci
14	993.5	15.3	1501	5	Aau76762 Plasmodiu
15	768	11.9	1115	2	AAR13457 Duffy rec
16	761.5	11.7	749	2	AAR70233 P. falcip
17	761.5	11.7	749	2	Aaw22479 Plasmodiu
18	761.5	11.7	749	3	AAY77901 P. falcip
19	754	11.6	1115	2	Aaw22478 Duffy ant
20	754	11.6	1115	3	AAY77899 P. vivax
21	750.5	11.6	972	5	Aau76761 Plasmodiu
22	743.5	11.5	1245	2	AAR70106 TNF-R-Pl.
23	704.5	10.9	1061	2	AAR70231 P. vivax
24	698	10.8	1028	2	AAR41044 Plasmodiu

25	684	10.6	350	5	ABB07656	Abb07656 P. falcip
26	655	10.1	778	2	AAR13456	Aar13456 Duffy rec
27	651	10.0	2459	8	ADO69869	Ado69869 Plasmodiu
28	637.5	9.8	3056	8	ADR43499	Adr43499 CSA-Bindi
29	601.5	9.3	3542	4	AAB62142	Aab62142 P. falcip
30	592	9.1	3147	9	AEb22180	Aeb22180 Codon opt
31	592	9.1	3553	9	AEb22174	Aeb22174 Plasmodiu
32	554.5	8.6	2994	9	AEb22176	Aeb22176 Plasmodiu
33	546.5	8.4	302	5	ABB07655	Abb07655 P. falcip
34	546.5	8.4	2913	2	AAW00384	Aaw00384 Plasmodiu
35	543.5	8.4	3346	9	AEb22178	Aeb22178 Plasmodiu
36	498.5	7.7	3060	2	AAW22475	Aaw22475 Plasmodiu
37	498.5	7.7	3060	3	RAY77905	Aay77905 Plasmodiu
38	495.5	7.6	2703	2	AAR70236	Aar70236 P. falcip
39	495.5	7.6	2710	2	AAW22482	Aaw22482 Plasmodiu
40	495.5	7.6	2710	3	RAY77904	Aay77904 P. falcip
41	492.5	7.6	697	9	ADZ72259	Adz72259 Plasmodiu
42	432	6.7	1507	3	AAB24128	Aab24128 Plasmodiu
43	414	6.4	762	9	ADZ72267	Adz72267 Plasmodiu
44	410.5	6.3	2485	3	AAB18172	Aab18172 Plasmodiu
45	409.5	6.3	2647	8	ADP25446	Adp25446 Plasmodiu

ALIGNMENTS

RESULT 1  
ABG73547

ID ABG73547 standard; protein; 1210 AA.

XX AC ABG73547;

XX DT 24-FEB-2003 (first entry)

XX DE P. falciparum BAEBL protein.

XX KW BAEBL; erythrocyte binding protein; protozoacide; immunostimulant;

XX KW malaria; parasite; vaccine; chromosome 13.

XX OS Plasmodium falciparum.

XX FH Key Location/Qualifiers

FT Misc-difference 185

FT /note= "A polymorphism resulting in variation from Val to Ile at this position is specifically claimed in claim 8"

FT Misc-difference 239

FT /note= "A polymorphism resulting in variation from Ser to Asn at this position is specifically claimed in claim 8"

FT Misc-difference 261

FT /note= "A polymorphism resulting in variation from Thr to Arg at this position is specifically claimed in claim 8"

FT Misc-difference 285

FT /note= "A polymorphism resulting in variation from Lys to Glu at this position is specifically claimed in claim 8"

XX WO200278603-A2.

XX PD 10-OCT-2002.

XX PP 29-MAR-2002; 2002WO-US010071.

XX PR 02-APR-2001; 2001US-0281130P.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Mayer G, Miller LH;

XX XX WPI; 2003-092869/08.

XX DR N-PSDB; ABG73547.

XX XX New vaccine against malaria Plasmodium falciparum parasite comprising Erythrocyte Binding Protein polypeptide.

PS	Claim 4; Page 51-54; 56pp; English.	
CC	This invention describes a novel vaccine composition comprising the	
CC	Plasmodium falciparum erythrocyte binding protein, BAEBL found on	
CC	chromosome 13. The composition is useful for preparing a medicament for	
CC	vaccinating a human against a malaria Plasmodium parasite and also has	
CC	protozoacide and immunostimulant activity. This sequence represents the	
CC	BAEBL polypeptide described in the disclosure of the invention	
XX		
SQ	Sequence 1210 AA;	
	Query Match 100.0%; Score 6481; DB 6; Length 1210;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MKGYFNIYFLIPIFLYNNVIRINESIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS 60	
Db	1 MKGYFNIYFLIPIFLYNNVIRINESIIGRTLYNRQDESSDISRVNSPELNNHKTNIYDS 60	
Qy	61 DYEDVNNKLSNFRVENSVKKGSLSLFINNKTYSYDIIPPSYRNDKFNLSLENEDNSG 120	
Db	61 DYEDVNNKLSNFRVENSVKKGSLSLFINNKTYSYDIIPPSYRNDKFNLSLENEDNSG 120	
Qy	121 NTSNFPANTSEISIGKDNQYTFIQKRTHLFACGIKRSIKWICRENSKITTVCVPDRK 180	
Db	121 NTSNFPANTSEISIGKDNQYTFIQKRTHLFACGIKRSIKWICRENSKITTVCVPDRK 180	
Qy	181 IQLCVANFLNRLTMEKFEIPLISVNTBAKLLYNKNEGKPSIFCNELRNSFSFPRSS 240	
Db	181 IQLCVANFLNRLTMEKFEIPLISVNTBAKLLYNKNEGKPSIFCNELRNSFSFPRSS 240	
Qy	241 FIGDDMDPGTDRVKGIYINTKSDYYKEQVKEKLANIKKEWKEKKNANLWNHMIVNHKG 300	
Db	241 FIGDDMDPGTDRVKGIYINTKSDYYKEQVKEKLANIKKEWKEKKNANLWNHMIVNHKG 300	
Qy	301 NISKECAIIPAEPOINLWIKWENENFLMEKRLFLNIKDKCVENKYYEACFGGCRLPSCS 360	
Db	301 NISKECAIIPAEPOINLWIKWENENFLMEKRLFLNIKDKCVENKYYEACFGGCRLPSCS 360	
Qy	361 SYTSFMKSKTQMEVLNLYKKKNSGVDKGNFLNDLFPKNNKNDLDDFPKNEKEYDDLCD 420	
Db	361 SYTSFMKSKTQMEVLNLYKKKNSGVDKGNFLNDLFPKNNKNDLDDFPKNEKEYDDLCD 420	
Qy	421 CRTATTIISKPLNGPAKNDVDIASQINVDLRGFGCNYKSNKSNWCTGTFTNKPGTC 480	
Db	421 CRTATTIISKPLNGPAKNDVDIASQINVDLRGFGCNYKSNKSNWCTGTFTNKPGTC 480	
Qy	481 EPPRRQTLCLGRYTLHRGHEEDYKEHLGASIEAQLLYKKEKEDENALCSIIQNSYA 540	
Db	481 EPPRRQTLCLGRYTLHRGHEEDYKEHLGASIEAQLLYKKEKEDENALCSIIQNSYA 540	
Qy	541 DLADIIKGSIIIDYIYKQKMEENLNKYNKDKGRNEESLKITPREKWDENKENYKVMASV 600	
Db	541 DLADIIKGSIIIDYIYKQKMEENLNKYNKDKGRNEESLKITPREKWDENKENYKVMASV 600	
Qy	601 LKNETCKDYKQKIPQFLRWPKWGGDRCEKREKEKIYSPFSFKVECKKCDENTCKN 660	
Db	601 LKNETCKDYKQKIPQFLRWPKWGGDRCEKREKEKIYSPFSFKVECKKCDENTCKN 660	
Qy	661 KCSEYKKWIDLKSEYKQVDKYTKDNKKMYDNIDEVKNKEANVYLKEKSKCKDVNFD 720	
Db	661 KCSEYKKWIDLKSEYKQVDKYTKDNKKMYDNIDEVKNKEANVYLKEKSKCKDVNFD 720	
Qy	721 DKIFNESPNEYDMCKKCDIBIKYLNIEIKYPTKHDIYDIDTFSDTFDGTGTPISINANINE 780	
Db	721 DKIFNESPNEYDMCKKCDIBIKYLNIEIKYPTKHDIYDIDTFSDTFDGTGTPISINANINE 780	
Qy	781 QOSGKDTNNTGNETSDSPVSHPEPESDAAINVEKLSGDESSSETRGLDINDFSVTNNVN 840	
Db	781 QOSGKDTNNTGNETSDSPVSHPEPESDAAINVEKLSGDESSSETRGLDINDFSVTNNVN 840	
Qy	841 EVHDAASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSGNEQSDNQNSHSDNSG 900	

841

EVHDAASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSGNEQSDNQNSHSDNSG

900

901

SLTIGQVPSEDNTQNTYDSQNPHRDTPNALASIPSDDKINEIEGPDSSRDSSENGRDTTS

960

901

SLTIGQVPSEDNTQNTYDSQNPHRDTPNALASIPSDDKINEIEGPDSSRDSSENGRDTTS

960

961

NTHDVVRTNIVSERRVNSHDFIRNGMANNNAHHQYITQIENNGIIRGOESAGNSVNYKD

1020

961

NTHDVVRTNIVSERRVNSHDFIRNGMANNNAHHQYITQIENNGIIRGOESAGNSVNYKD

1020

1021

NPKRSNFSENDHKNIQEVNSRDTKRVREIILKSKQKCNNEYSMEYCTYSDERNSSP

1080

1021

NPKRSNFSENDHKNIQEVNSRDTKRVREIILKSKQKCNNEYSMEYCTYSDERNSSP

1080

1081

GPCSREERKLCQISDYCLKYFNFYSIEYNNCKSEIKSEIKSPYKCFKSEQSSIPYPAAG

1140

1081

GPCSREERKLCQISDYCLKYFNFYSIEYNNCKSEIKSEIKSPYKCFKSEQSSIPYPAAG

1140

1141

GILVWIVLLSSASRMKSNSEYVDIGESNIEATFEENNYLNKLSRIFNQEVQETNISDYS

1200

1141

GILVWIVLLSSASRMKSNSEYVDIGESNIEATFEENNYLNKLSRIFNQEVQETNISDYS

1200

1201

EYNYNEKNMY 1210

1201

EYNYNEKNMY 1210

RESULT 2

AAU76759

ID

AAU76759 standard; protein; 1143 AA.

XX

AAU76759;

XX

DT

21-MAY-2002 (first entry)

XX

Plasmodium falciparum erythrocyte binding protein BBP2 protein sequence.

XX

Erythrocyte binding protein; BBP; malaria parasite; protozoacide;

KW

vaccine; immune response inducer; Plasmodium falciparum merozoite; SABP;

KW

salicylic acid binding protein; EBA-175.

XX

Plasmodium falciparum.

XX

PH

Key Location/Qualifiers

FT

Region 147..762

FT

/label= BBP2 RII

FT

/note= "Region II of BBP2 protein"

XX

WO200211756-A2.

XX

14-FEB-2002.

XX

07-AUG-2001; 2001WO-US024725.

PR

07-AUG-2000; 2000US-0223525P.

XX

(ENTR-) ENTREMED INC.

PA

PI

Narum DL, Sim KL;

XX

WPI; 2002-227117/28.

XX

PT

Paralog of erythrocyte binding protein-175 polypeptide sequence useful

PT

for inducing immune response to Plasmodium falciparum merozoites, that

PT

causes malaria, in a patient, and for diagnostic and prognostic purposes.

XX

Claim 2; Fig 2; 52pp; English.

PS

XX

The present invention relates to a new polypeptide comprising a paralog

CC

of erythrocyte binding protein-175 (EBA-175) polypeptide sequence. The

CC

invention is useful for inducing an immune response to Plasmodium

CC

falciparum merozoites in a patient. The method of the invention comprises

CC

administration of isolated SABP (salicylic acid binding protein) binding

Db	847	EVDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSGNEQSDNQENSHSSDMSG	906
Qy	901	SLTIGQVPSEDNTQNTYDSQNPHRDTPNALASLPDDKINEIEGFPDSSRDSENGRGDITTS	960
Db	907	SLTIGQVPSEDNTQNTYDSQNPHRDTPNALASLPDDKINEIEGFPDSSRDSENGRGDITTS	966
Qy	961	NTHDVRRTIVSERRVNSHDFIRNGMANNAAHQYITQIENNGIIRGQESAGNSVNYKD	1020
Db	967	NTHDVRRTIVSERRVNSHDFIRNGMANNAAHQYITQIENNGIIRGQESAGNSVNYKD	1026

1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494</
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

```

0001 1081 GPCSRERKRLCLQISDILCLKIPNFISIRHINCIKSEIKSPERKCFKSEQSSS 1134
0002 1087 GPCSRERKRLCCQISDYLKVFNFYSIEYINCISEIKSPYKCFKSEQSSM 1140
0003
0004
0005
0006
0007
0008
0009
0010
0011
0012
0013
0014
0015
0016
0017
0018
0019
0020
0021
0022
0023
0024
0025
0026
0027
0028
0029
0030
0031
0032
0033
0034
0035
0036
0037
0038
0039
0040
0041
0042
0043
0044
0045
0046
0047
0048
0049
0050
0051
0052
0053
0054
0055
0056
0057
0058
0059
0060
0061
0062
0063
0064
0065
0066
0067
0068
0069
0070
0071
0072
0073
0074
0075
0076
0077
0078
0079
0080
0081
0082
0083
0084
0085
0086
0087
0088
0089
0090
0091
0092
0093
0094
0095
0096
0097
0098
0099
0100
0101
0102
0103
0104
0105
0106
0107
0108
0109
0110
0111
0112
0113
0114
0115
0116
0117
0118
0119
0120
0121
0122
0123
0124
0125
0126
0127
0128
0129
0130
0131
0132
0133
0134
0135
0136
0137
0138
0139
0140
0141
0142
0143
0144
0145
0146
0147
0148
0149
0150
0151
0152
0153
0154
0155
0156
0157
0158
0159
0160
0161
0162
0163
0164
0165
0166
0167
0168
0169
0170
0171
0172
0173
0174
0175
0176
0177
0178
0179
0180
0181
0182
0183
0184
0185
0186
0187
0188
0189
0190
0191
0192
0193
0194
0195
0196
0197
0198
0199
0200
0201
0202
0203
0204
0205
0206
0207
0208
0209
0210
0211
0212
0213
0214
0215
0216
0217
0218
0219
0220
0221
0222
0223
0224
0225
0226
0227
0228
0229
0230
0231
0232
0233
0234
0235
0236
0237
0238
0239
0240
0241
0242
0243
0244
0245
0246
0247
0248
0249
0250
0251
0252
0253
0254
0255
0256
0257
0258
0259
0260
0261
0262
0263
0264
0265
0266
0267
0268
0269
0270
0271
0272
0273
0274
0275
0276
0277
0278
0279
0280
0281
0282
0283
0284
0285
0286
0287
0288
0289
0290
0291
0292
0293
0294
0295
0296
0297
0298
0299
0300
0301
0302
0303
0304
0305
0306
0307
0308
0309
0310
0311
0312
0313
0314
0315
0316
0317
0318
0319
0320
0321
0322
0323
0324
0325
0326
0327
0328
0329
0330
0331
0332
0333
0334
0335
0336
0337
0338
0339
0340
0341
0342
0343
0344
0345
0346
0347
0348
0349
0350
0351
0352
0353
0354
0355
0356
0357
0358
0359
0360
0361
0362
0363
0364
0365
0366
0367
0368
0369
0370
0371
0372
0373
0374
0375
0376
0377
0378
0379
0380
0381
0382
0383
0384
0385
0386
0387
0388
0389
0390
0391
0392
0393
0394
0395
0396
0397
0398
0399
0400
0401
0402
0403
0404
0405
0406
0407
0408
0409
0410
0411
0412
0413
0414
0415
0416
0417
0418
0419
0420
0421
0422
0423
0424
0425
0426
0427
0428
0429
0430
0431
0432
0433
0434
0435
0436
0437
0438
0439
0440
0441
0442
0443
0444
0445
0446
0447
0448
0449
0450
0451
0452
0453
0454
0455
0456
0457
0458
0459
0460
0461
0462
0463
0464
0465
0466
0467
0468
0469
0470
0471
0472
0473
0474
0475
0476
0477
0478
0479
0480
0481
0482
0483
0484
0485
0486
0487
0488
0489
0490
0491
0492
0493
0494
0495
0496
0497
0498
0499
0500
0501
0502
0503
0504
0505
0506
0507
0508
0509
0510
0511
0512
0513
0514
0515
0516
0517
0518
0519
0520
0521
0522
0523
0524
0525
0526
0527
0528
0529
0530
0531
0532
0533
0534
0535
0536
0537
0538
0539
0540
0541
0542
0543
0544
0545
0546
0547
0548
0549
0550
0551
0552
0553
0554
0555
0556
0557
0558
0559
0560
0561
0562
0563
0564
0565
0566
0567
0568
0569
0570
0571
0572
0573
0574
0575
0576
0577
0578
0579
0580
0581
0582
0583
0584
0585
0586
0587
0588
0589
0590
0591
0592
0593
0594
0595
0596
0597
0598
0599
0600
0601
0602
0603
0604
0605
0606
0607
0608
0609
0610
0611
0612
0613
0614
0615
0616
0617
0618
0619
0620
0621
0622
0623
0624
0625
0626
0627
0628
0629
0630
0631
0632
0633
0634
0635
0636
0637
0638
0639
0640
0641
0642
0643
0644
0645
0646
0647
0648
0649
0650
0651
0652
0653
0654
0655
0656
0657
0658
0659
0660
0661
0662
0663
0664
0665
0666
0667
0668
0669
0670
0671
0672
0673
0674
0675
0676
0677
0678
0679
0680
0681
0682
0683
0684
0685
0686
0687
0688
0689
0690
0691
0692
0693
0694
0695
0696
0697
0698
0699
0700
0701
0702
0703
0704
0705
0706
0707
0708
0709
0710
0711
0712
0713
0714
0715
0716
0717
0718
0719
0720
0721
0722
0723
0724
0725
0726
0727
0728
0729
0730
0731
0732
0733
0734
0735
0736
0737
0738
0739
0740
0741
0742
0743
0744
0745
0746
0747
0748
0749
0750
0751
0752
0753
0754
0755
0756
0757
0758
0759
0760
0761
0762
0763
0764
0765
0766
0767
0768
0769
0770
0771
0772
0773
0774
0775
0776
0777
0778
0779
0780
0781
0782
0783
0784
0785
0786
0787
0788
0789
0790
0791
0792
0793
0794
0795
0796
0797
0798
0799
0800
0801
0802
0803
```

Query Match 24.0%; Score 1557; DB 2; Length 1435;  
Best Local Similarity 28.5%; Pred. No. 2e-88;  
Matches 410; Conservative 214; Mismatches 467; Indels 348; Gaps 42;







WPI; 1995-115452/15.

New hybrid peptide(s) for binding cytokine(s) - comprising a malaria parasite peptide capable of binding a red blood cell and a receptor peptide.

Example A; Page 48-51; 93pp; English.

Hybrid peptides for binding cytokines, comprising a malaria parasite (Plasmodium falciparum) peptide (capable of binding to a red blood cell (RBC)) and a receptor peptide are claimed. AAR70103-25 are examples of these hybrid peptides. AAR70105 is a fusion of tumour necrosis factor receptor (in accordance with H Loetscher et al Cell, Vol. 61, 351-359) and erythrocyte binding antigen (EBA) (175 kDa). The use of cytokine receptors not normally found on RBCs means that the cytokine can bind harmlessly to the RBC without deleterious effect. The RBC protects the hybrid peptides from excretion from the kidney, and due to steric hindrance prevents the cytokines binding to a receptor in another cell. Glycophorin binding peptide (GBP) 130 or GBPH (GBP homologue) are the pref. malaria parasite peptides used, others include EBA 175, PMMSA (pre major merozoite surface antigen) and the Duffy binding receptor molecule (eg. exhibited by Plasmodium vivax). These peptides bind to pref. glycophorin A, B and C, sialo glycoproteins, found on the surface of RBCs. The hybrid peptides are thus used to lower the levels of free cytokines in the circulation to reduce pathological damage. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)

Sequence 1604 AA;

Query Match 23.8%; Score 1543.5; DB 2; Length 1604;

Best Local Similarity 28.4%; Pred. No. 1.6e-87;

Matches 405; Conservative 215; Mismatches 459; Indels 345; Gaps 42;

18 NVIRINESIGRTLYNQDESS--DISRVNSPELANNHKNTIYSDYEDVNNKLNSFVZ 75  
 180 NVGTEDSGTARNEYDIDKENEKFLDVYKFPNELDKKKYGNVQKTD-----KKFTFIE 233  
 76 NKSVKKRSISFINN-----KTKSY---DIIPPSYSYRNDKFNLSLSEDSNGTNSNNF 127  
 234 NK-----LDILNNSKFNKWKSYGTPDNDIKNWSLIN-----KINNEFMNNYQSP 280  
 128 ANTSEISIGKDNKO-----YTFIQKRTH-----LFCAGIKRKSIRKWC 165  
 281 LSTS--SLIKQNKVYPINAVRVSRILSFLDSRINNGRNTSSNNEVLSNCKRKGKMWDC 338  
 166 RENSEKIT--VCVPRKQLCVANFLNSRLETMEKFEFLISVNTAEKLYNKNEGKOPS 224  
 339 KCKNDRSNYYCIPDRRIQLCIVNLIIITKTYTETMDQHFTEASKESQILLKKNDNKYNS 398  
 225 IFCNELNSFDSFRSSFIGDDMDFGNVDKVGINTKFSDYKKNVKEKLNNIKKEMWE 284  
 399 KFCMDLKNISFLDYGLHANGMDMPGGYSTKAENKIQEVFGAHGEISEHKIKNFKKEWN 458  
 285 KKNANLNMHIMVNHKGNISKECAIIPAEFPQINLWIKENFENFLMEKKRLFLNKKRCVE 344  
 459 EFREKLWEAMLSSEKNNIN--NCKNIPOEELQITQWIKWEGHEFLLRDRNSKLPKSKCN 517  
 345 NKYTEACFGGRLPCSSYTSFMKSKTKTOMEVLTHLYKKNSGVDDKNNFLNDFK--KNKN 403  
 518 NTLYEACEKECIDPCMYRDIWIIIRSKFEPWHTLSKEYETQK--VPKNAENYLIKISENN 575  
 404 D--LDDFFKN--EKYDLDCCRYTATIKSPNGPAKNDVDIASQINVDLGRGQNVKS 460  
 576 DAKVSLILNNDABYSKYCDCKHTTILVKSVLNGNDNTIKEKREHIDLDDFSFGCDKNS 635  
 461 --NNEKSNWCTGTETNKPFGTCEPRRTQLCLGRTYLLHRGHEEDYKEHLGASIYBAQL 518  
 636 VDTNKTWECKNPYLSTKDCVCPPEQLCLGNIDRIYDKNLLMIKEHILAIYESRI 695  
 519 LKYYKEDNALCISITONSADLADIIKGSDDIIKQYKGMENLKNVNDKKRNEESL 578  
 696 LKRYKKNDDKEVKCIINKTFADIRDIIGGTDVWNLDSNRKLVGKINTNSKYVHRNKND 755

QY 579 KIFREKWDENKENVWVMSAVLKNKCKYDKFQKIPOFLRWFKWGGDDFCRKRKEKI 638  
 DB 756 KLFREDEWKKVKKDVMNVISWFKDKTVCKE--DDIENIPQFFRWFSEWGGDDYCDQTKMI 814  
 QY 639 YSPESFKVECKKDCDENTCNKCKSEYKWKIDLKSKSEYKQV---DKYTKDNKKQMYDNI 695  
 DB 815 ---ETLKECKEPCEDNCKSKNSYKWKISKKKEEYNKQAKYQYQYQKNNYKMY--- 868  
 QY 696 DEVKNEANYVLKESKCKEDVNPDDKI FNESNEYEDMCKCDEIKVL----- 744  
 DB 869 SEFKSIPEVILKYSEKSNLNFDEFKELHSDYKKNCTMCPVKDVPISIIIRNEQT 928  
 QY 745 -----NE----- 746  
 DB 929 SQRAVPEENTEIAHRTETPSISBPGKNEQKERRDDSLSKISVSPENSRPETDAKTSNL 988  
 QY 747 -----IKYPT-----KHIYDID----- 760  
 DB 989 LKLKGDVDISMPRAVIGSSPNNDINVTQGDNIGVNSKPLSDVDVDPKKELEDQNSDES 1048  
 QY 761 -----TFSDTFGDTPLSI----- 774  
 DB 1049 EETVNVHISKSPSINNGDDSGSGSATVSESSSNTGLSIDDRNGDTFVQTQDTANTEDV 1108  
 QY 775 ---NANINQQSGKDTNTGNETSDSP---VSHEPESDAAINVEKL-----SGDESSE 823  
 DB 1109 IRKENADKDEKGADEERHSTESLSLSPPEKMLTONEGNSLNHEEVKHTNSDNDVQ 1168  
 QY 824 TRGILDIN-----DPSVTNNVNEVH-----DASNTQGSVNTSDITN- 860  
 DB 1169 SGGIVNVMVEKELKDTLENPSSSLDEGKAHEELSEPNLSDDQDMSNTPGFLDNTSEETTE 1228  
 QY 861 --GHSESLNR-----TTNAODIKIGRSGNEQSDN-----QENSSSHSSDNGSLTI-GQ 906  
 DB 1229 RISNNEYKVNREDEKTLTKYEDIVLKSHMNRSDDGELYDENSJLSTVNDESEDAEK 1288  
 QY 907 VPSEDNTQNTY-----DSQNPHRDTPNALA-----SLP-----SDDKINE----- 941  
 DB 1289 MKGNDTSEMNSHNSQIHESDQKNDKMTGDLGTHVQNEISVPVTGCEIDKLESKESK 1348  
 QY 942 -----IBGPDSSRDSNGRGTTSNTHDVR-----RTNIVSERRVNSHDFI 982  
 DB 1349 IHKAEERLSHTDIHKINPEDRNSNTLHLKDIRNEENERHLTNQINISQERDLQKHGF- 1407  
 QY 983 RNMANNNAHQVITQ-----IENNGIIRGOESAGNSVNYKON-----PKESNESSE 1030  
 DB 1408 ---HTWNHLHGDGVSESRQINSHHGNRQDRGNSGVNLRNNSNNNNFNIPSYNL--- 1461  
 QY 1031 NDHKNIQENYNSRDTKVRREEIKLSKONCKNNEYSMEYCTYSDERNSSPGPCSREERKK 1090  
 DB 1462 YDKLDDLLENRNDSTTKELIKLAELKNEISVKYCDHMHIEIPLKCTCKEKTNRN 1521  
 QY 1091 LCCOISDYCLKYFNFYISYIYNNCIBKSEIKSPYKCFKSEGOSSI 1134  
 DB 1522 LCCAVSDYCMYFTYDSEYYNCTKGFDDPSYTCFRKEAFSSM 1565

RESULT 7  
 AAR41043  
 ID AAR41043 standard; protein; 1786 AA.  
 XX AC AAR41043;  
 XX AC AAR41043;  
 DT 24-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 22-MAR-1994 (first entry)  
 XX CD4-EBA175 fusion protein.  
 KW Merozoite; Erythrocyte Binding Antigen 175; malaria; HIV; env;  
 KW human immunodeficiency virus; envelope glycoprotein; hybrid protein;  
 KW red blood cell; erythrocyte; AIDS; molecular machine.





Db 1105 TKEVEDVLKSHNRSDGELDYDENSGLTVNDESDAEAKKNGDTSEMSHNSQHIE 1164  
 Qy 918 -DSQPHRDTPNALA-----SLP-----SDDKINE-----REGFDSRDS 951  
 Db 1165 SDOQKNDMKVTGDLGTHVQNEISVPTGTGEIDKLRSEKSKIHKAEEERLSHTDIHKIN 1224  
 Qy 952 ENRGDGTTSNTHDVR-----RTNIVSERRVNSHDFIRNGMANNNAHQYITQ-- 998  
 Db 1225 PEDRNSWTLHLKDIRNEERHLTNQINISQERDLQKHGF-----HTMNNLHGDGVRSERS 1280  
 Qy 999 ---IENNGLIRGOESAGNSVNYKDN-----PKRSNFSSSENDHKKNIQEYNSRDTKRV 1048  
 Db 1281 QINHSHHGNGQDRGNSGNVLNWSNNNNFNIPSRYNL-----YDKKLDLDLYENRNDSTT 1337  
 Qy 1049 REBIIKLSKQNKNNVEYMYCTYSDBRNSSPGCSREERKLCQOISDYCLKYFNFYSI 1108  
 Db 1338 KELIKKLAEINKCENEISVKYCDHMIHEEIPLTKCTKEKTRNLCVAVSDYCMSTFTYDSE 1397  
 Qy 1109 EYNYCIKSEIKSPYKCFKSEGO 1132  
 Db 1398 EYNYCTKREFDPSYTCFRKEAFS 1421

# RESULT 9 ABR82498

ID ABR82498 standard; protein; 616 AA.

XX AC ABR82499;

XX 20-NOV-2003 (first entry)

DT Synthetic erythrocyte-binding protein EBA-175 RII sequence.

XX BFP; erythrocyte-binding protein; malaria; pathogen; pProtozoasidae;  
 KW immunostimulant; vaccine; EBA-75; EBA-175 RII.

XX Synthetic.

OS Plasmodium falciparum.

XX WO2003062374-A2.

XX 31-JUL-2003.

XX 12-NOV-2002; 2002WO-US036368.

XX 09-NOV-2001; 2001US-0345051P.

XX (ENTR-) ENTREMED INC.

PI Narum DL, Liang H, Fuhrmann S, Sim K;

XX WPI; 2003-598746/56.

DR N-PSDB; ACF35993.

XX New synthetic genes encoding erythrocyte-binding protein of Plasmodium  
 PT falciparum, useful as immunization agents for inducing anti-malarial  
 PT response and/or treating malaria, or for raising antibodies for passive  
 PT immunization.

XX Claim 5; Fig 9; 73pp; English.

XX The invention relates to new synthetic nucleotide sequences (I) encoding  
 CC erythrocyte-binding protein (EBP) of a malaria pathogen, where codon  
 CC usage of the synthetic nucleotide sequence is altered compared to a  
 CC naturally occurring sequence of the EBP to approximate codon usage of a  
 CC host of the malaria pathogen. (I) and the purified erythrocyte binding  
 CC protein are useful for the production of a medicament for induction of an  
 CC anti-malarial immune response in a human or animal. The methods are  
 CC useful for inducing an anti-malarial immune response in a host of the  
 CC malarial pathogen, optimizing expression in P. pastoris of a recombinant  
 CC Plasmodium falciparum region II EBA-75 erythrocyte-binding protein, and  
 CC obtaining the pure protein, and obtaining a DNA vaccine for inducing an

CC anti-malarial immune response. (I) is useful in the production of anti-  
 CC malaria DNA vaccines with improved immunogenicity, in expression vectors  
 CC for the production of malaria EBPs, or as an immunization agent for  
 CC inducing anti-malarial response and/or treating malaria. The EBPs are  
 CC useful for raising antibodies or serum for passive immunization and  
 CC treatment of malarial diseases. The present sequence represents a EBA-175  
 CC RII synthetic protein without N-glycosylation sites

XX Sequence 616 AA;

Query Match 18.5%; Score 1200.5; DB 6; Length 616;  
 Best Local Similarity 38.4%; Pred. No. 1.5e-66;  
 Matches 246; Conservative 111; Mismatches 236; Indels 47; Gaps 13;

Qy 154 CGIKRSIKIKWCRENSEKIT-VCPDRKIQLCVANFLNSLETMEKFEIPLISVNTSEAK 212  
 Db 14 CREKRGMKWDCKKNDKRSNYVICPDRRIQLCIVLWLAIIKTYTETKTMKOHFIEASKESQ 73  
 Qy 213 LLYNKHEGKDPISFCNELRNSFSDPRSSPFGDDMDFGGNTDRVKGYNITKFSYDYEKNV 272  
 Db 74 LLLKKNDNKYNSKFCNDLKNSEFLDYGHLAWGNDMDFGGYSTKAENKIQEVFKGAHGEISE 133  
 Qy 273 EKLNNIKKEWKEKNKANLWNHMIWNHGNISKECAIIPAEPOINLWIKWENNENFLMEKK 332  
 Db 134 HKIKNFRKKWNEFREKLMWELSEKNNIN-NCKNIPOEELQITQWIKWHGEFLLERD 192  
 Qy 333 RLFLNITKDKVENKYEACFGGRLPCSSYTPSMKSKTQMEVLTNLYKKKNSGVDKKNF 392  
 Db 193 NRAKLPKSKCNALYEACEKICIDPCMKYRDWIRSKFEWHTLSKEYETQK--VPKENA 250  
 Qy 393 LNDLFK-KNNKND--LDDFFKN-EKEYDDLCDCRYTATIIKSPNGPAKNDVDIASQINV 448  
 Db 251 ENYLKISENNDKAVSLLNNDCAEYSKYCDCKHTTLVKSVLNGNDNTIKKREHIDL 310  
 Qy 449 NDLRGFGCNVKS--NNEKSNWCTGTFTNKPFGTCEPPROTLCGLRTYLLHRGHEEDYE 506  
 Db 311 DDFSKFGCDKNSVDNTKWECKPKYKLTQKVCVPRQELCLGNIDRIYDKNLLMIKE 370  
 Qy 507 HLGASIIYBAQLKYYKKEKDENALCSIIQNSYADLADIIGKSDIIKDYVKGKHEENLAK 566  
 Db 371 HILAIAYESRILKRYKNDKDEKCKIIQKTADIRDIIGGTDYWNDSLNRKLVGKINT 430  
 Qy 567 VNKDKRNEESLKIPEKWDENKENVKMSAVLNKKECTCKDYDFQKTPQLRPFKEW 626  
 Db 431 NSNYVHRNKQNDKLFDEWVKVIKDVWNVISWFKDKTVCKE-DDIENIPQFRWFSEW 489  
 Qy 627 GDDFCEKRKEIYSPSFKVECKKDCDENTCKNCKSEYKKWIDLKSEYKQV---DKY 683  
 Db 490 GDDYCDQKTKMI---ETLKECKEKCPCDDNCKRCKNSYKEWISKKEETNKQAKQOEY 546  
 Qy 684 TKDKNKKMYDNIDEVKNKEANVYLKESKCEKDVNFDDKIFNESPNEYEDMKCKDEIKY 743  
 Db 547 QKGNNYKMY---SEPKSIKPEVYLKYSKCSNLNFEDEPKELHSDYKKNCTMCPVK- 602  
 Qy 744 LNEIKYPTKTHYDIIDTFSDTFGDTGTPISINANINEQOS 783  
 Db 603 -----DVPISIIIRN-NEQTS 616

## RESULT 10 ABR82498

ID ABR82498 standard; protein; 616 AA.

XX AC ABR82498;

XX 20-NOV-2003 (first entry)

DE Synthetic erythrocyte-binding protein EBA-175 RII.

XX BFP; erythrocyte-binding protein; malaria; pathogen; pProtozoasidae;  
 KW immunostimulant; vaccine; EBA-75; EBA-175 RII.

OS Synthetic.







Db 643 NNGSLQDQTKQVCTNMNVYTRKLAIEIQSVKY--DKDKLP----SLAKDKNVTTF 697  
Qy 707 LKEKSKCKDVNPDKIFNBNPEYEDMKCKD-----EIKYLNIEIKYPTKHDIYDI--- 759  
Db 698 LKENAKNCNIDF-TKIFDQDLKFKERCSCMDQVLEKVKNEKMLSDSNSDSDATDISEK 756  
Qy 760 -----DTPFS 763  
Db 757 NGBEELYNHNSVSVASGNKEIEKSKDEKQPEKAKQTNGLTVRTDKDSRNKGDAT 816  
Qy 764 DT----- 765  
Db 817 DTGNSPENLKVQHGHTNGETIKPEPPKLPESSEFLQSQEQLAEAAQKQEEBPKKQEE 876  
Qy 766 -----FGDGTPIISINANINEQSQG 785  
Db 877 EPKKKQEEQKQEQKQEQEERQKQEQEQIQDQSQGLDQSSKGVASEQNEISSGQ 936  
Qy 786 D-----TNTGNSSETSDSPVSH-EPESDAAIN--VEKLSGDESSSETRGI 827  
Db 937 EQNVKSSPEVVPQETTSNGSSQ--DTKISSTEPNENSVVDRATDSMNLDPKVNHNEM 994  
Qy 828 LDIN-----DPSVTNNVNEVDA-----SNTQ-----GVSNTS 856  
Db 995 SDPNTPEPDAKLKDKKQVDDAKKELQSTVSRIESNEQDVQSTPPEDTPTVEGKGDKA 1054  
Qy 857 DI-----TNGHSESLN-----RTTN----- 872  
Db 1055 EMLTSPHATDNSESESLNPTDDIKTTDGVVKEQELGGESATETSKNLEKPKDVEPS 1114  
Qy 873 ---AQDIKIGRSGNEQSD-----NOENSS-----HSDNDSGLTIGQ 906  
Db 1115 HEISEPLGSLVTGKESELLKSKSIETKGTDPKSDQEDATDDVNSRDDNNLS--- 1171  
Qy 907 VPSEDNTQNTYDSNP-----HRDTPNALASLPDDDKINIEGPDSS 948  
Db 1172 -NSVDNQSNVLNREDPIASETEVVSPEPDSRSRIMTTEVPTTVK-PDEKRSVEVGEKA 1229  
Qy 949 RD-----SENGRG-DTTSNTHDVRRTNIVSER-----RVNSHDFFIR 983  
Db 1230 KEIKVPEVPVRAIGPEMENSVSQSPNVEDVEKETLISENGLHNDTHRGNISSEKDLID 1289  
Qy 984 NGMANNAHGYITQIENNG-IIRQESAGN-----SVNYKDNPKRSFSE----- 1030  
Db 1290 IHLRNEAGSTLDDRRNGEMTEGSESDVGELQEHNFSTQQDKEDFQDIASDREKEEI 1349  
Qy 1031 -----NDH-----KKNIOEYNSRDTKRVREI 1054  
Db 1350 QKLNINIGHEDEDLKMDRTEDSDGNSHLYNNLSSEKMEQYNNRDASKDREILN 1409  
Qy 1055 LSKQKCNNEYSMEYCTYSDERNSSPGCSREERKKLCCQISDYCLKYFNYFIEYVNCI 1114  
Db 1410 RSNNTNCSNEHSLKYCYMERNKDLLETCTSEDKRLHLCCBISDYCLKFFPKSIETVDC 1469  
Qy 1115 KSEIKSPYCKFKSEGOSSPYFAAGLIVVILLSSASRMGKSNEEYDIGESNI-EAT 1173  
Db 1470 QKEPDDPTNCFKQRKRTSMHYIAGGGIALLUPLGASVYRKNLDDKDEKGYDSNLDSA 1529  
Qy 1174 FE-ENNYLNKLS-RIFNQEQVETNISDYSEYNNYKKNY 1210  
Db 1530 FEYNNKYKLPYMVVDQINNVNSDLYSEGIYDDTTTF 1568

## RESULT 14

AAU76762  
ID AAU76762 standard; protein; 1501 AA.  
XX AAU76762;  
AC AAU76762;  
XX AAU76762;  
XX 21-MAY-2002 (first entry)  
XX Plasmodium falciparum erythrocyte binding protein EBP5 protein sequence.

KW Erythrocyte binding protein; EBP; malaria parasite; protozoacide;  
KW vaccine; immune response inducer; Plasmodium falciparum merozoite; SABB;  
XX salicylic acid binding protein; EBA-175.

OS Plasmodium falciparum.

XX WO200211756-A2.

XX 14-FEB-2002.

XX 07-AUG-2001; 2001WO-US024725.

XX 07-AUG-2000; 2000US-0223525P.

XX (ENTR-) ENTREMED INC.

XX Narum DL, Sim KL;

XX WPI; 2002-227117/28.

XX Paralog of erythrocyte binding protein-175 polypeptide sequence useful  
PT for inducing immune response to Plasmodium falciparum merozoites, that  
causes malaria, in a patient, and for diagnostic and prognostic purposes.

XX Example 9; Fig 2; 52pp; English.

CC The present invention relates to a new polypeptide comprising a paralog  
of erythrocyte binding protein-175 (EBA-175) polypeptide sequence. The  
invention is useful for inducing an immune response to Plasmodium  
falciparum merozoites in a patient. The method of the invention comprises  
administration of isolated SABB (salicylic acid binding protein) binding  
domain polypeptide. The invention is further useful for research  
applications, vaccine candidates, blocking peptides, diagnostics and  
prognostics, where diseases of Plasmodium falciparum include malaria. The  
polypeptide of the invention and antibodies are useful for diagnosis of  
malaria, and for detection of P. falciparum in culture media and in  
biological samples such as biological tissues and fluids. The present  
amino acid sequence represents the Plasmodium falciparum EBP5 protein  
which is one of several (AAU76759-AAU75762) EBA-175 paralogs

XX Sequence 1501 AA;

Query Match 15.3%; Score 993.5; DB 5; Length 1501;  
Best Local Similarity 21.7%; Pred. No. 5.1e-53;  
Matches 343; Conservative 228; Mismatches 472; Indels 537; Gaps 63;

Qy 1 MKGVFNI-----YFLIPLIFLYNVIRINESIIGRTLYNRQDESSDLSRVNSPELNHNKT 55  
Db 3 MKGKMMCLFFFSIILYVVLCTYVLGISEYL-----KERPQGLNVETNNNNNNNNNS 57

Qy 56 NIYDSYEDVNNKLINSFVEN-KSVKKGRSLSFNNKTSYDIIPFSYSYRNDKFNLSLSE 114

Db 58 N--SNDAMSFVNEVIR-FIENEKDDKEDKKVKIISRPVEN-----TLHRYVPVSSFLNIKK 109

Qy 115 NEDNSGNTSNPNFANTSEISIGKDNQYTFIQKTHLFCAGIKRKRKIKWICRENSKITV 174

Db 110 YGRKGTYLARNSPV-----QRSYIRGCKGRKSTHTWIC-ENKGNINI 150

Qy 175 CVPDRKIQLCVA--NFLNSRLTMEK--FKEIFLSVNTAEKLLYNKNGKDPSPICNE 229

Db 151 CIPDRVQLCITALQDLKNSSGSETTDKRLDRKVFDSAMYETDLLWNKYGFDFDFDCCD 210

Qy 230 LRNSFDFRSSFIGDDMDFGNTDRVKGYINTKFSYDYKKNYKVEKNLNKIKKEWKNKAN 289

Db 211 VKNSYLDYKDVIFGTDLT---KNNISKLEVEBSLKEFFK-KDSSVLN--PTAWRRYGR 263

Qy 290 LNNHMI--VNHKGNISKECAIPAEFPOINLWKEH---NENFLMEKKRLFLNIKOCVCE 344

Db 264 LMKMTIQPYAHLG-----CRKPDENEPNIRNWILEWGYKNCRLMKEKEKL---LTGEC 315

Qy 345 NKKYACFGGCRLLPCSSYTSFMKSKTOMEVLTNLKCK-----KNSGVYDKNNFLNDLF 397

Db 316 NRKSDCSGTCNNECTYRSNLNRQYEVSIILGKKIKVRYTIFPRKIVQPNAL-DFL 374





Db 30 CRSQPEFNK-----IPERNVQI-----HISNIFKEYKENNVDIIFGTLNYYNN 74  
 Qy 418 LC-DCRYTATTIIFSLFNGPAKNDVDIASQINVDLGRFGCNYK--SNNEKSNWC--TGFTFN 474  
 Db 75 FCKEPELVAANKYLNKAPNAKS--PRIYKSEHESSVFGCKTKISKYKKNWCYNNKVT 134  
 Qy 475 KPGTCPPRRQTLCLGRTYLLHRGHEEDYKEHLGASIIYEAQLLKYYKEKDNALCSI 534  
 Db 135 KPEGVCGPPRRQQLCLGYIFLIRDGNEBGLKDHINKAANYEAMHLKKEYNAGDKICNA 194  
 Qy 535 IONSADYADLADIIGSDIIDKYGKMBEENLVK-----NKDKRNEESLKIIFREKWDEN 589  
 Db 195 ILGSYADIGDIVRGDLVWRDINTNKLSEKFOKIFMGGNSRKKQNDNN---ERNKWNKQ 251  
 Qy 590 KENVKVMASVLNKNKCTKDYDFKQIPQFLRWFKEWGDGDFCHKKEKIYSFESFKVECK 649  
 Db 252 RNLWSSVMVXHIPGKTCRKHNNFEKIPQFLRWLKEWGDGDFCEEMGTVEVKLEKI---CE 308  
 Qy 650 KDCDENTCKNCKSEYKKNWIDLKSEYEKQVDKTKOK--NKK--MYDNDIDEVKNKANV 705  
 Db 309 NKNCSEKCKNACSSYEKWKERKNEYNLQSKFDSKLNKKNLYNKPEDSK-----A 363  
 Qy 706 YLKEKSECKDVNPDCKIFNESPNYEYDMCKKCD---EIKYLNKIK---YPTKHDYDI 759  
 Db 364 YLRSEKQCSNIEFNDETFT--FPNKYKEACMVCENPSSSKALKPIKTNVPIESKSEL 422  
 Qy 760 DTFSDFGDTPTPISINANINEQQSGK--DTSNTGNSETSDSPVSHPESDAAINV-----812  
 Db 423 SSLTDSKNTNSGGNGYGRQISKRDDVHHDGPKVK--SGEKEVPKIDAAVKTENEFT 481  
 Qy 813 -----EKLSDGESSETRGILDINDPS--VTNNVNEVHDASNTQGSVNSDITNG 861  
 Db 482 SNRNDIBGKSKGDSHPVSHSKDIKNEEPQVVSLENLPKIEEKWESSDSPTIT-----535  
 Qy 862 HSSESLARTNAQD-----IKIGRSGNEQS-----DNQENSHSDNSSLTIGQVPS 909  
 Db 536 HIEAKQGSNSNDPAVSGRESKQVNLHTSRIKENEEGVIKTDSDSKSIBSKIPS 595  
 Qy 910 EDN-----TQNTYDSQNPHRDTPNALASLPDDK--INEIGFDSRSDSENGRGTTSNT 962  
 Db 596 DQNNHSDLSQANEDSQNKETINP-----PSTEKNLKEIHYKTSDDSDHGSKISIEP 651  
 Qy 963 HDVARTNIVSERR-----VNSHDFIRNGMANNNAHYYITQIENNGIIRGOE 1009  
 Db 652 KELTEBPLTDKTESAAIGDKNHESVKSADIPQSEIHNSDNRDRIVSE-----SVW---Q 704  
 Qy 1010 ESAGNSVNYK-----DNPKRSNFSSENHKKNIQYNSR 1043  
 Db 705 DSSGSSMSTESIRTDN---KDFKTSEDIAPSINGRNSR 739

RESULT 17  
 AAW22479  
 XX AC AAW22479 standard; protein; 749 AA.  
 XX AC AAW22479;  
 XX DT 07-OCT-1997 (first entry)  
 XX DE Plasmodium ebl-1.  
 XX KW DBL gene family; SAPP; sialic acid binding protein; vaccine; therapy;  
 XX KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;  
 XX KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;  
 XX KW Plasmodium.  
 XX OS Plasmodium falciparum.  
 XX PN WO9640766-A2.  
 XX PD 19-DEC-1996.  
 XX PF 07-JUN-1996; 96WO-US009508.

XX 07-JUN-1995; 95US-00487826.  
 PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA Sim KL, Chitnis C, Miller LH, Peterson DS, Su X, Wellems TE;  
 FI WPI; 1997-052231/05.  
 DR N-PSDB; AAT72894.  
 XX New malaria vaccines - contains cysteine-rich DBL family protein binding  
 PT domains homologous domains of the Duffy and sialic acid binding proteins.  
 XX Disclosure; Page 41-42; 96pp; English.  
 XX This sequence represents ebl-1 of Plasmodium. Ebl-1 belongs to the Duffy  
 CC binding like (DBL) family of genes which have homology to the Duffy  
 CC antigen binding protein (DABP) and sialic acid binding protein (SAPP)  
 CC conserved regions (see AAT72889 and AAT72888 respectively). The var  
 CC family of genes modulate cytoadherence and antigenic variation of  
 CC Plasmodium infected erythrocytes. SAPP and the Duffy antigen binding  
 CC protein (DABP) are soluble proteins that appear in the culture  
 CC supernatant after infected erythrocytes release merozoites. DABP and SAPP  
 CC mediate the binding of merozoites and schizonts to the erythrocyte  
 CC surface. These proteins are necessary for erythrocyte invasion by the  
 CC parasite. This sequence can be used in the compositions of the invention.  
 CC The compositions are for the treatment and prevention of malaria, and  
 CC comprise either a nucleotide sequence or encoded polypeptide of the var-  
 CC 1, var-2, var-3 or var-7 genes of the DBL gene family, a family of genes  
 CC having homology with conserved regions of DABP and SAPP. The compositions  
 CC are used for the treatment and prevention of malaria. They are also used  
 CC in the preparation of vaccines for inducing a protective immune response  
 CC in a mammal to Plasmodium merozoites (especially Plasmodium falciparum or  
 CC Plasmodium vivax)  
 XX Sequence 749 AA;  
 SQ

Query Match 11.7%; Score 761.5; DB 2; Length 749;  
 Best Local Similarity 27.3%; Pred. No. 7.2e-39;  
 Matches 207; Conservative 148; Mismatches 282; Indels 121; Gaps 31;

Qy 359 CSYTSFMKSKTKQMEVLNLYKKNSGVKNFLNDFKNNKNNDLDDFPKN-EKEYDD 417  
 Db 30 CRSQPEFNK-----IPERNVQI-----HISNIFKEYKENNVDIIFGTLNYYNN 74  
 Qy 418 LC-DCRYTATTIIFSLFNGPAKNDVDIASQINVDLGRFGCNYK--SNNEKSNWC--TGFTFN 474  
 Db 75 FCKEPELVAANKYLNKAPNAKS--PRIYKSEHESSVFGCKTKISKYKKNWCYNNKVT 134  
 Qy 475 KPGTCPPRRQTLCLGRTYLLHRGHEEDYKEHLGASIIYEAQLLKYYKEKDNALCSI 534  
 Db 135 KPEGVCGPPRRQQLCLGYIFLIRDGNEBGLKDHINKAANYEAMHLKKEYNAGDKICNA 194  
 Qy 535 IONSADYADLADIIGSDIIDKYGKMBEENLVK-----NKDKRNEESLKIIFREKWDEN 589  
 Db 195 ILGSYADIGDIVRGDLVWRDINTNKLSEKFOKIFMGGNSRKKQNDNN---ERNKWNKQ 251  
 Qy 590 KENVKVMASVLNKNKCTKDYDFKQIPQFLRWFKEWGDGDFCHKKEKIYSFESFKVECK 649  
 Db 252 RNLWSSVMVXHIPGKTCRKHNNFEKIPQFLRWLKEWGDGDFCEEMGTVEVKLEKI---CE 308  
 Qy 650 KDCDENTCKNCKSEYKKNWIDLKSEYEKQVDKTKOK--NKK--MYDNDIDEVKNKANV 705  
 Db 309 NKNCSEKCKNACSSYEKWKERKNEYNLQSKFDSKLNKKNLYNKPEDSK-----A 363  
 Qy 706 YLKEKSECKDVNPDCKIFNESPNYEYDMCKKCD---EIKYLNKIK---YPTKHDYDI 759  
 Db 364 YLRSEKQCSNIEFNDETFT--FPNKYKEACMVCENPSSSKALKPIKTNVPIESKSEL 422  
 Qy 760 DTFSDFGDTPTPISINANINEQQSGK--DTSNTGNSETSDSPVSHPESDAAINV-----812  
 Db 423 SSLTDSKNTNSGGNGYGRQISKRDDVHHDGPKVK--SGEKEVPKIDAAVKTENEFT 481

```
QY 813 -----EKLSDGESSESTRGILDINDPS--VTNNVNEVHDASNTQGSVNTSDITNG 861
Db 482 SNRNDIEGKSGKGDHSPVHSDKDKNBPQVVSENLPKIEKQWESSDSIPIT----- 535
QY 862 HSESSLNRTTNAOQ-----IKIGRSGNEQS-----DNOENSSHSDNDSGLTIGOVPS 909
Db 536 HIEAEKGQSSNSDNDPAVVGRESKDVNLHTSERIKENEGVKTDDSSKSIEISKIPS 595
QY 910 EDN-----TON-TYDSQNPHRDTPNALASIPSDOK-INEIEGFDSSRDSNGRGDTTSNT 962
Db 596 DQNNHSDLSQANEDSNQGNKETINP-----PSTEKNLKEIHYKTSDDSDHSGSIKSEIEP 651
QY 963 HDVVRTNIVSERR-----VNSHDFIRNGMANNNAHHOYITQIENNGIIRGOE 1009
Db 652 KELTESPLTDKKTESAAIGDKNHESVKSADIFQSEIHNSDNRDRIVSE-----SVW---Q 704
QY 1010 ESAGNSVNYK-----DNPGRSNFSSNDHKNIQVNSR 1043
Db 705 DSSGSSMSTESIRTDN---KDFKTSEDIAPSINGRNSR 739
RESULT 18
AAY77901
ID AAY77901 standard; protein; 749 AA.
AC AAY77901;
XX
DT 13-JUN-2000 (first entry)
DE
DE P. falciparum ebl-1 polypeptide.
XX
KW DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
KW DABP; Sialic Acid Binding Protein; SABP; malaria; vaccine; immunisation;
KW protozoacide.
XX
OS Plasmodium falciparum.
XX
FN US5993827-A.
XX
PD 30-NOV-1999.
XX
PF 07-JUN-1995; 95US-00487826.
XX
PR 10-SEP-1993; 93US-00119677.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Sim KL, Chitnis C, Peterson DS, Su X, Wellem's TE, Miller LH;
XX WPI; 2000-194198/17.
XX DR N-PSDB; AAZ98283.
XX
Isolated protein binding domains from Plasmodium vivax and Plasmodium
falciparum erythrocyte binding proteins useful for vaccinating against
malaria.
XX
Claim 7; Col 53-58; 93pp; English.
XX
The invention relates to ebl-1 polypeptides that are encoded by the DBL
(Duffy-binding like) gene family. The ebl-1 proteins are substantially
identical to the Duffy Antigen Binding Protein (DABP) and sialic acid
Binding Protein (SABP), which are soluble proteins that appear in the
culture supernatant after erythrocytes infected with malaria release
merozoites. Immunohistochemical studies indicate that DABP and SABP are the
respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
used to vaccinate against malaria, especially caused by P. falciparum.
CC Immunization with the polypeptide provides effective protection against
CC malaria. The present sequence represents the ebl-1 polypeptide
XX
SQ Sequence 749 AA;
```

Query Match 11.7%; Score 761.5; DB 3; Length 749;

```
Best Local Similarity 27.3%; Pred. No. 7.2e-39;
Matches 207; Conservative 148; Mismatches 282; Indels 121; Gaps 31;
QY 359 CSYTSFMKSKTKQMEVLNLYKKKSGVDKNNFLNDFKNNKNDLDDPFKN-EKEYDD 417
Db 30 CRSGPFPNK-----IFPERVQI-----HISNIFKEYKENVDIIFGLTNEYNN 74
QY 418 LC-DCRYTATITIKSFLNGPAKNDVDIASQINVDNLRGFGCNYK-SNNEKSWNC-TGTFTN 474
Db 75 FCXEKPELVSAAYNLKAPNAKSPRIYKSEHESSVFGCKTKISKVKKKWCYNNKVT 134
QY 475 KPGTCCEPRRQTLCLGRTYLLHRGHEEDYKEHLLGASIEAQLLYKYKEKENALCSI 534
Db 135 KPEGVCGPFRROQLCLGYIFLRDGNBGLKDHINKAANYEAHMLKEKYEAGDKTCNA 194
QY 535 IQNSYADLADIIGSDIIOYGYKMBEENLKV-----NKDKRNEESLKIPEKMWDEN 589
Db 195 ILGSYADIGDIVRGLDVWRDINTNKLSEKFKIFMGGSNRKKQNDNN---ERNKWWEKQ 251
QY 590 KENVWKMVAVLKNETCKDYDKFKQIPQPLRWFKEGDDFCRKRKEKIYSPFSFVCK 649
Db 252 RNLIWSMVKHIPKGTCKRHNNFEKIPQPLRLWKEGDFCEMGTEVQLEKI---CE 308
QY 650 KKDCEKCKNCKSEYKMWIDLKSEYKQVDKYTKK--NKK--MYDNIDEVKNKEANV 705
Db 309 NKNCSKKCKNACSSYERKWKERKNYNLQSKKFDSDKLNKNLKNKPFDSK-----A 363
QY 706 YLKEKSKCKDVNFDKIFNESPNEDYEMCKKCD---EIKYLNKIK---YPKTKHDIYDI 759
Db 364 YLRESKQCSNIFNDETF-T-FNKYKEACWVCENPSSSKALPKITNVPPIESKSEL 422
QY 760 DTFSDTFGDTPIISINANINEQSGK--DTSNTGNSSETSDSPVSHPEPSDAAINV----- 812
Db 423 SSLTDKSKNTPNSSGGNGYDGRQISKRDVVDHDPKEVK-SGEKEVPKIDAAVKTENEFT 481
QY 813 -----EKLSDGESSESTRGILDINDPS--VTNNVNEVHDASNTQGSVNTSDITNG 861
Db 482 SNRNDIEGKSGKGDHSPVHSDKDKNBPQVVSENLPKIEKQWESSDSIPIT----- 535
QY 862 HSESSLNRTTNAOQ-----IKIGRSGNEQS-----DNOENSSHSDNDSGLTIGOVPS 909
Db 536 HIEAEKGQSSNSDNDPAVVGRESKDVNLHTSERIKENEGVKTDDSSKSIEISKIPS 595
QY 910 EDN-----TON-TYDSQNPHRDTPNALASIPSDOK-INEIEGFDSSRDSNGRGDTTSNT 962
Db 596 DQNNHSDLSQANEDSNQGNKETINP-----PSTEKNLKEIHYKTSDDSDHSGSIKSEIEP 651
QY 963 HDVVRTNIVSERR-----VNSHDFIRNGMANNNAHHOYITQIENNGIIRGOE 1009
Db 652 KELTESPLTDKKTESAAIGDKNHESVKSADIFQSEIHNSDNRDRIVSE-----SVW---Q 704
QY 1010 ESAGNSVNYK-----DNPGRSNFSSNDHKNIQVNSR 1043
Db 705 DSSGSSMSTESIRTDN---KDFKTSEDIAPSINGRNSR 739
RESULT 19
AAW22478
ID AAW22478 standard; protein; 1115 AA.
XX
AC AAW22478;
XX
DT 12-SEP-1997 (first entry)
DE Duffy antigen binding protein.
XX
KW DBL gene family; SABP; sialic acid binding protein; merozoite; malaria;
KW Duffy antigen binding protein; DABP; erythrocyte; var-1; var-2; var-3;
KW var-7; vaccine; therapy; immune response; Plasmodium.
XX
OS Plasmodium vivax.
XX
PH Key Location/Qualifiers
```

```

FT Misc-difference 1065. .1067
FT /note= "encoded by CTGCTGTGTTA"
FT
FT WO9640766-A2.
FT
FT 19-DEC-1996.
FT
FT 07-JUN-1996; 96WO-U009508.
FT
FT 07-JUN-1995; 95US-00487826.
FT
FT (USSH ) US DEPT HEALTH & HUMAN SERVICES.
FT
FT Sim KL, Chitnis C, Miller LH, Peterson DS, Su X, Wellens TE;
FT
FT WPI; 1997-052231/05.
FT
FT N-PSDB; AAT72889.
FT
FT New malaria vaccines - contains cysteine-rich DBL family protein binding
FT domains homologous domains of the Duffy and sialic acid binding proteins.
FT
FT Example 1; Page 33-35; 96pp; English.
FT
CC This sequence represents the full length Duffy antigen binding protein
CC (DABP). DABP and the sialic acid binding protein (SABP) are soluble
CC proteins that appear in the culture supernatant after infected
CC erythrocytes release merozoites. DABP and SABP mediate the binding of
CC merozoites and schizonts to the erythrocyte surface. These proteins are
CC necessary for erythrocyte invasion by the parasite. This sequence can be
CC used in the compositions of the invention. The compositions are for the
CC treatment and prevention of malaria, and comprise either a nucleotide
CC sequence or encoded polypeptide of the var-1, var-2, var-3 or var-7 genes
CC of the DBL gene family, a family of genes having homology with conserved
CC regions of DABP and SABP. The compositions are used for the treatment and
CC prevention of malaria. They are also used in the preparation of vaccines
CC for inducing a protective immune response in a mammal to Plasmodium
CC merozoites (especially Plasmodium falciparum or Plasmodium vivax)
FT
FT SQ Sequence 1115 AA;
FT
FT Query Match 11.6%; Score 754; DB 2; Length 1115;
FT Best Local Similarity 24.5%; Pred. No. 3.7e-38;
FT Matches 296; Conservative 177; Mismatches 416; Indels 318; Gaps 57;
FT
FT 199 FKEIFLSVNTAEKLLYNKNGKDPSTFCNELRNSFSDFRSSFFGDDWDFGNTDVKYGY 258
FT 22 YKDDFSIT-----LNYHEGKKYLILKR-----KLEKANNRDVCNFF 59
FT
FT 259 IN-TKFSDDYKERNVEKLNNIKKEWBE-KKANLWNEHIVNHNKGNISKECAIIPAEFPQI 316
FT 60 LHPFQVNNVLLERTIETLLCEKNEYVKGNGYKL-----AKGH---HC-----VEEDNL 105
FT
FT 317 NLWIKENWENFLMEKKRLFLNIHKCVENKYEACFGGRLPCS-SYTSFPMKSKTOMEV 375
FT 106 ERWLOQTNER-----RSEENIKYK--YGVTELKIKYQAQNGKRSSRIILKES 149
FT
FT 376 L-----TNLYKKKSGVDKNNFLNDLPKNNKNDLDDFPKNEKEDDLCDCRY----- 423
FT 150 IYGAHNFGNSYMEKGDDGDKTG-----EEKDGEHKTDSKTDNGKANNVLMDDYETSSN 204
FT
FT 424 -----TATIIFKSLFNGPAKN-----DVD-----IASQINVNDLRGF-----G 455
FT
FT 205 GQAPGTLNDLVLEFVTGHEGSRKNSNGNPYDIDHKHTISSAI-IN-HAFLQNTWKN 261
FT
FT 456 CNYK-SNNKSWNCTGFTTNKPGTCBPPTQTLCLGRITVLLHRGHEEDY-----KEH 507
FT 262 CNYKRRERDWC-----NTKDVCI-PDRRYQLCMKELTNLVNNTDTHFRDITPRKLY 316
FT
FT 508 LLGASIVFA-----QLLYKYEKEDENALCSIIQNSVADLADIIGSDIINKDYGKK 559
FT 317 LKRKLIDYAAVEGDLKLANNRY-NKD---FCKDIRWSLGDGFDGIIMGTMEGIGYSKV 372
FT
FT 560 MEENLNKV-NKDKRNEESLKI PREKWNDEKENKENVKMSAVLNK-----ETCKDYDK 612
FT

```

---

```

Db 373 VENNLSIFGTDEKAQR-----RKQWNNESSAQIWTAMVSVKRLKGNFIWICKLNV 427
Qy 613 FQKIPQLRWFKEWGDFFCEKREKIKYSPFVKVC-----KKQCDENTCKNKSEY 665
Db 428 VNIPEQIYRWRWGRDYSELTEV---OKLKEKCDGKINYTDKKYCKVPCPCNACKSY 484
Qy 666 KKWIDLKSEYEKQVDKYTKDKNKNKYDNIDEVKNKEANV-----YLKESKCKECKVNF 720
Db 485 DQMITRKNQMDVLSNKFISVKNAB-----KVQTAGIVTPVDYILKQELDEFENVA 536
Qy 721 DKTFNESPNEYEDMCKKCDIKYLNEIKYKPKTHDIYDIDTFSDTFG-DGTPI----- 772
Db 537 NEI-NKRDGAYIELC-----VCSVEEAK-KNTQEVTVTVNDAAKSOATNSNPISQ 589
Qy 773 -----SINANINEQOSGKDTNCTNSETSP-----SPVSH-----PSDAIN 812
Db 590 KAEKVPGDSTHGNV---SGQDSSTTGKAVTGGQNGNTPAESDVQORSIDAESVSAN 646
Qy 813 E-----KLSGD-----ESSSETRGILDINDPSPVNNVNEVHDASNTQGSV----- 852
Db 647 DPQKSVSKRSDDTASVTGIAEACKENLGASNSRPSSESTVEANSPGDDTVNSASIP 706
Qy 853 -----SNTSDITNGHSESSLNRTTNAQDIKIGRSG-----NEQSDNQNSHSS 896
Db 707 NPLVTPYNGLRHKSNDSDGPAESMANPDSNK---GETGKQDNDMAKATKDSNSS 762
Qy 897 DNGSLT-----IGQVPSEDNTQNTYDSQNPHRDTPNALASLPSDDKINEIG 944
Db 763 DGTSSATGDTTDAVDREINKGVPEDRDKTVGSKDGGEDNSANKDAATVVGEDR 822
Qy 945 FDSRD-----SENG-----RGDTTSN---THDVRRTIV 971
Db 823 GGTNDRSKNDTEKNGASTPDSKQSEDATAALSKESTESGDRTYNDTNSLENKGG 882
Qy 972 SERRVNSHDFIRGMANNNAHQYIT-----QIENNGIIRQOESAGNSVNYKDNPK 1023
Db 883 KEKDLQKDFKSNPTNEEPNSDQTTDAEGHDRDSIKND---KAERKHKMKDFTTKNTN 939
Qy 1024 RSNFSEN-----DHKKNIQENSRDTPKVRREIILKSKQKCNKNEYSNEYCTYSD 1080
Db 940 SHLNSNNLSNGKLDIKEYKYRDVATREDIILMSVRKCNNNISLEYCSNVEDKISS- 998
Qy 1081 GPCSRERERKLCQOISDYCLKYFNFYNIYVNCIKSEIKSPYKCPKSEG-QSSIPYFAA 1139
Db 999 NTSREKSKMLCCSIDFCLNYFDVYSEYLSCKKEFEPSYKCTFGKFKDKTYFAA 1058
Qy 1140 GGILWVIVILLSSASRMKSNBEYDIGESNIEATPEE-NNYLNKLSRI-----FNQ 1193
Db 1059 AGAL--LILLIASRKMKNDS-----EATFNEFEYCDNIHRIPLMPNNIEHQ 1108
Qy 1194 TNISDYS 1200
Db 1109 STPLDYS 1115

```

---

```

RESULT 20
AA777899
ID AA777899 standard; protein; 1115 AA.
XX
XX AA777899;
XX
XX DT 13-JUN-2000 (first entry)
XX
XX DE DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
XX KW DABP; Sialic Acid Binding Protein; SABP; malaria; vaccine; immunisation;
XX KW protozoa.
XX OS Plasmodium vivax.
XX
XX PN US5993827-A.

```





CC (RBC)) and a receptor peptide are claimed. AAR70103-25 are examples of  
 CC these hybrid peptides. AAR70106 is a fusion of tumour necrosis factor  
 CC receptor (in accordance with H Loetscher et al Cell, Vol. 61, 351-359)  
 CC and Plasmodium vivax Duffy binding receptor. The use of cytokine receptors  
 CC not normally found on RBCs means that the cytokine can bind harmlessly to  
 CC the RBC without deleterious effect. The RBC protects the hybrid peptides  
 CC from excretion from the kidney, and due to steric hindrance prevents the  
 CC cytokines binding to a receptor in another cell. Glycophorin binding  
 CC protein (GBP) 130 or GBP Homologue (H) are the pref. malaria parasite  
 CC peptides used, others include EBA 175 (175 kDa erythrocyte binding  
 CC antigen), Pwmsa (pre major merozoite surface antigen) and the Duffy  
 CC binding receptor molecule (eg. exhibited by Plasmodium vivax). These  
 CC peptides bind to pref. glycophorin A, B and C. sialo glycoproteins, found  
 CC on the surface of RBCs. The hybrid peptides are thus used to lower the  
 CC levels of free cytokines in the circulation to reduce pathological  
 CC damage. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-  
 CC 2003 to standardise OS field)  
 XX  
 SQ Sequence 1245 AA;

Query Match 11.5%; Score 743.5; DB 2; Length 1245;  
 Best Local Similarity 23.1%; Pred. No. 2e-37;  
 Matches 303; Conservative 192; Mismatches 466; Indels 353; Gaps 59;

Qy 96 DIIPPSYSYRNDKENSISENEDSGNTNSNFPANTSEISIGKONKQYTFIQKTHL---F 152  
 Db 20 DSVCPQKYLHPQNSISCTCKCHKGTLYNDPCPGQDTPCRCESGSFTASENHLRHCL 79  
 Qy 153 ACGIKRKSIAKWI-----CREN-----SEKITVCPDRKIQLCVANFLNSR 192  
 Db 80 SCSCRKEMQVEIISCTVDRTVCGCRKQYRHYWSENLPQCF---NCSLCLNGTVH-- 134  
 Qy 193 LETWEKPEKIFLISVNTAEKLLYNNKGDPSIFCNELRNSF----- 234  
 Db 135 LSCOEKQNTV---CTCHAGPFLRENE---CVSCSNCKKSECTKLCPLQIENVKGTED 186  
 Qy 235 -----SDFRSSFIG-----DDMDPGGNTDRVKGYIN-TKFSDDYKKNVKEKLN 277  
 Db 187 SGTKDDFSITILNYHEGKKYLIILKRKLEKANNRDVNCFFLHFSQVNNVLLERTIETLE 246  
 Qy 278 IKKEW--WEKNKANLNMHMIWNHKGNIKSECAIIPAEPOINLWIKENENFLMEKKRLF 335  
 Db 247 CKNEYVGKNGKYLAKGHCV-----EEDNLERWLQGTNER----- 282  
 Qy 336 LNTKDKVENKKYKACFGGRLPCS-SYTSFMKKSQTQMEVL-----TNLYKKKNSGV 387  
 Db 283 -----RSEENIKYK--YGVTELKIKYAQMNGKRKRIRLKEISYGAHNFNGSGYMEGKDG 335  
 Qy 388 DKNFLNDLPKONKNDLDOPFKNEKEYDDLCDCRY-----TATIIKSLANGPAK 437  
 Db 336 DKTG-----BEKGEHKTDSKTDNGKANLMLDYETSSNGOPAGFLDNVLFVVTGHEG 390  
 Qy 438 N-----DVD-----IASQINVDLGRF-----GCNYK-SNNEKSNWCTCTFT 473  
 Db 391 NSRKNSSNGGNPYDIDHKTKISSAI-IN-HAFLONTVMKNCVYKREKREDWDC----- 442  
 Qy 474 NKPGTCEPPRQTLCLGRVTLVLRHGEEDY-----KXHLIGASIYEA-----QL 518  
 Db 443 NTKKDVCIPDRRYQLCKMELTNLVNNTDTNFRDITFRKLYLKRKLIYDAAVEGDILLKL 502  
 Qy 519 LKYKYEKDENALCSIIQNSVADLADIIKGSDDIIKDYVGKMEENLKV-NKDKRNEES 577  
 Db 503 NTRYI-NKD---PCKDRLWSLGDPGDIIIMGTD-MEIGYSKVVENNLSIFGTDEKAQOR- 556  
 Qy 578 LKIPREKWMDEKENVKWSAVLKNK-----ETCKDYDKFKIPQFLRFKFWGDDFC 631  
 Db 557 -----RKQWNNESKAQIWTAMMYSYKELKGNFIWICKLVAVNIEPQIYRWIRWGRDYV 612  
 Qy 632 EKRKEKJYSFESFKEC-----KKKDCDENTCNKCSYKWIIDLKKESEYKQVDKYT 684  
 Db 613 SELPTEV---QKLEKCDGKINYTDKKVKVPPQCNACKSYDQWITRKQKQWDLVLSNKF 669  
 Qy 685 KDNKHKQYDNIDIVKNEANV-----YLKKEKSECKDVNFDDKIFNESPNYEYDMCKCD 739

Db 670 SVKNAE-----KVQTAGIVTPYDILKQELDEFNEVAFENEI-NKRDGAYTELCL 716  
 Qy 740 EIKYLNEIKYPTKHDIDYDIDTFPSDTFG-DGTPI-----SINANINEQQS 783  
 Db 717 -VCSVEEAK-KWTEVVTNVNNAKSOAINSPISQPVDSKAEKVPGDSTHGNV---S 771  
 Qy 784 GKDTNVTGNETSD-----SPVSH-----PESDAAINVE-----KLSGD----- 818  
 Db 772 GQDSSTTGKAVTGQNGNQTPAESDVQSRSDIAESVSANKVDPQKSVKSRSDTASVTGI 831  
 Qy 819 -ESSSETRGLDINDSVNNVNEVHDSANTQGSV-----SNTSDIT 859  
 Db 832 AEAGKENLGASRRPSESTVEANSPGDDTVNSASIPVSGENPLVTPYNGLRHSKONSDS 891  
 Qy 860 NGHSESLARTTNAQDIKIGRSG-----NEQSDNQENSSHSSDNGSGSLT----- 903  
 Db 892 DGPAESMANPDSNSK-----GETGKGQDNMAKATKDSNSSDGTSATGDTTDAVDREIN 947  
 Qy 904 IGVQPSEDNTQNTYDSQNPHRDTPNALASLPDDKINEIEGFDSSRD-----SENG--- 954  
 Db 948 KGVPEDRDKTVSGKGGEDNSANKDAATVVGEDRIRENSAGGSTNDRSKNDTEKNGAST 1007  
 Qy 955 -----RGDTTSN---THDVRTNIVSERRVNSHDFIRNGMANN 990  
 Db 1008 PDSKQSEDATALKTESLSTESGDRRTTNDTTNSLENKONGKEKDLQKHDFKSDNTPEE 1067  
 Qy 991 AHQYIT-----QIENNGIIRGOESAGNSVNYKDNPKRNSFSEN---DHKKNIQE 1039  
 Db 1068 PMSDQTTDAEGHDDRSIKND---KAERRKHMNKDTFTKNTNSHLSNNLSNGLDKIKE 1124  
 Qy 1040 YNSRDTKRVREBIIKLSKQNKCNNEYSMEYCTVYSDERNSSPGPCSREERKLCQCSIDYC 1099  
 Db 1125 YKTRDVKATREDIILMSSVRKCNNNISLBYCNSVEDKISS-NTCSREKSKNLCSSISDFC 1183  
 Qy 1100 LKYFNEYSIEYNYCIKSEIKSPYKCFKSEG-QSSIPYFAAGILVIVILLSS 1152  
 Db 1184 LNYFDVYSIEYLSCKMKEFDPYKCTKGGFKIDKITYFAAAGALLILLIAS 1237  
 RESULT 23  
 AAR70231  
 ID AAR70231 standard; protein; 1061 AA.  
 XX  
 AC AAR70231;  
 DT 25-MAR-2003 (revised)  
 DT 22-SEP-1995 (first entry)  
 XX  
 DE P. vivax DABP.  
 XX  
 KW DABP; duffy antigen binding protein; binding domain; merozoite; malaria;  
 KW therapy; vaccine.  
 XX  
 OS Plasmodium vivax.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..325  
 FT /label= Binding\_domain  
 XX  
 PN WO9507353-A2.  
 XX  
 PD 16-MAR-1995.  
 XX  
 XX 07-SEP-1994; 94WO-US010230.  
 XX  
 PR 10-SEP-1993; 93US-00119677.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Sim KL, Chitnis C, Miller LH, Peterson DS, Su X, Welles TE;  
 XX WPI; 1995-123427/16.  
 DR







Db 393 DGEKADPOKLEV--KQQTDDKRSQSLGPHHTDERATLGETHWEKDTETAGGSTL----- 446  
Qy 926 TPNALASLPD-----DKINE-----IBGFSSRDSSENGRGDTTSNTHDVRRITVSE 973  
Db 447 TPEQNVSVASDNGNVPCKGKNEGATGALSABSLKSNESVHKITDNTHTHGLEKNGGNE 506  
Qy 974 RRVNSHDFIRNGMANNA----- 991  
Db 507 KDFQKHDFMNDMLNDQASDHTSSDQTSDDHTSSDQTSDDHTSSDQTSDDHTSSDQTS 566  
Qy 992 -----HHOYITQIENNGIIRGOESAGNSVNYKDNPKRSNFFSSNDHKNIQIENSR 1043  
Db 567 DQITDTEGHR--DNVRNPEIKSEDSKGDPMNSNSNELYSHNNLNRRKLNRDQYEHR 624  
Qy 1044 DTRVREEIIKSKQKNNNEYSMEYCTYSDERNSSPGCSREBRKKLCCQIISDYCLKYP 1103  
Db 625 DVKATREKIILMSVKNKNNRASVKYCNITIEDRLMS-STCSRRERKNLCCSISDFCLNYP 683  
Qy 1104 NFYSIEYNCIKSEIKSPYKCFKSEGOSSIPYPAAGGIILVIVILLSSASRMKMSNEEY 1163  
Db 684 ELXSYEFYNCMKFEFDPSECFTKGSGTGIVVPATGGAFLIILLFPASWNAASNDYEE- 742  
Qy 1164 DIGESNIEATFEE 1176  
Db 743 -----EATFDE 748  
RESULT 27  
ID ADO69969 standard; protein; 2459 AA.  
XX ADO69969;  
XX AC  
XX XX  
DT 29-JUL-2004 (first entry)  
XX Plasmodium falciparum virulence factor var O protein.  
DE antimalarial; vaccine; gene therapy; virulence factor; var O; Plasmodium;  
XX malaria.  
KW Plasmodium falciparum.  
XX OS  
XX WO2004037856-A2.  
XX PN  
XX PD 06-MAY-2004.  
XX PF 24-OCT-2003; 2003WO-EP013341.  
XX PR 25-OCT-2002; 2002CA-02409897.  
XX PA (INSP ) INST PASTEUR.  
XX PI Puijalon O, Le Scanf C, Lavergne A, Bentley G, Badaut C;  
PI Igonet S;  
XX DR WPI; 2004-399926/37.  
XX DR N-PSDB; ADO69968.  
XX PT New virulence factor var O polynucleotide of Plasmodium falciparum,  
XX useful in preparing a composition for treating or preventing malaria.  
XX PS Claim 9; SEQ ID NO 2; 84pp; English.  
XX CC The invention relates to an isolated or purified virulence factor var O  
CC polynucleotide comprising a nucleic acid sequence that is 65, 80 or 95%  
CC identical to a sequence comprising 7378 base pairs (bp) or its fragment.  
CC The polynucleotide is useful in preparing a composition for treating or  
CC preventing Plasmodium species related disease, particularly malaria. This  
CC sequence corresponds to the virulence factor varO protein.  
XX SQ Sequence 2459 AA;

Query Match 10.0%; Score 651; DB 8; Length 2459;  
Best Local Similarity 23.9%; Pred. No. 3.3e-31;  
Matches 245; Conservative 120; Mismatches 307; Indels 352; Gaps 43;  
Qy 75 ENKSVKKKRSLSF-----INNKTYSYDIIPPSYSYRNDKFNSLSNEDN-----SGNTN 123  
Db 1550 ENSSTQKDLPEAFDPCPKSIGRCNCPCPKPEPKYCVKTAHDIRKENGKGNFDINLKGNNV 1609  
Qy 124 SNN-----PANTSISIGKONKQYTFIQKRTHLFACGIRKRSI-- 161  
Db 1610 TYNDNCNAKREDYANQONGETCKFEVSSWSIGIINNENE-----STGRDFKIGE 1660  
Qy 162 KWICRENSE--KITVCVPDRKIQLCVANFLNSRLEWKEKEIFLISVNTAEKLLYN--- 216  
Db 1661 VMECKNETDGNKVKCVPPRRKDMC-----LKKLODIRVDDISDSSTLLKBIQE 1709  
Qy 217 --KNECKD-----PSIFCNE-----LRNSFSDFRSSFIGDD-----MDFGGNTDR 254  
Db 1710 VAKNGNDIIRNLPKYPCNEDVICMYKYSFADLGDIVRGTDKYKDVIGSISGNNSEQ 1769  
Qy 255 VKGYINTKFSDYK-----EKVKEKLNNIKGEWKEKNKANLNMHIMV-----HKGNISKE 305  
Db 1770 IBEENLKTIFENIQKTDENFQKYTNLELFRSAWMDANRDKIWKAMTCNAPDEAKIYITKE 1829  
Qy 306 CAII-----PAEE---PQINLMIKEWNEFLMEKKQLFLNIKDKC----- 342  
Db 1830 GGYISPITWTQNHCHGNDPPDYIYIPLRLWISSESYCLAQKD-FLETMKNCNCKK 1888  
Qy 343 -----VENKYEACFGGRLPCSSYTSFMKSKTOMBEVLTNLYK-----KONS 385  
Db 1889 KNDNTCEQTKYGAC-RDCKKCEEYKFKDWKAQFETQNKAYKAIYKNAUTSSGRHSN 1947  
Qy 386 GVDKN--NFLNDLFKKNKNDLDD-----PPKNEKEYD 416  
Db 1948 GIDEDIKFKVEKLEQNCQKNSVDADKYLEGGSVCRFRKFKVKTDTHEKNVAFHTPLSYK 2007  
Qy 417 DLCDCRYTATIKSFLNGPAKN-DVDIASQINVDLRGFG-----CNYK 459  
Db 2008 EHCEC-----AKNFDPLDECPVDNNECKYTGIGSCPEKFNHKKLEWTNYV 2053  
Qy 460 SNNKSNMCTGTFTNKFPCTCEPPRQTLCL-GRTYLLHR-GHEEDYKKEHLGASLVEAQ 517  
Db 2054 LNNK-----SNKNSAIVPPRRRQCLQNLTRNLSRLNKEKSKFEGILLISASEAK 2104  
Qy 518 LLKYKYKEKDNALCSIIQNSYADLADIIKGSIIKDYTGKKMEENLNK-VNKRKRNEE 576  
Db 2105 MLTEQYRENPAKAL-QAIKYSPADIGNIIGDDII-----GNVISVOLKLNGLNGKKNITS 2159  
Qy 577 SLKIFREKWDENKENVWVMSAV-----LKNKETCKDYKFKQKIPOFLRWFKEWGDFFCE 632  
Db 2160 TL-----WWEANKEKIWNAMMCLHWRKKTATSSPSHDNIDKEDQFLRWFQWGENFCA 2213  
Qy 633 KRKEIYSPESFKVECKKDC-----DENTCKNKSEYKWKWDLKSKSEYQVDKYTK 685  
Db 2214 TRKE-----LYEKLNNECKSVCEKASKGNVNLKCTKACEEYKSVLKKTYEIQKDKYDK 2270  
Qy 686 DKKNKMYNDIEVKNKEANYLYKESKECKDNVFDKIFNESP-NYEYD-----MCKK 737  
Db 2271 EFNKTL-----NNKNALEFL--NVQCISEYFSDSKWESPYDTFDDTLKGTVDCKK 2320  
Qy 738 CD----- 744  
Db 2321 HEPTPAIKPSKPASPEDKKLVDPDSLPIPIQPOPSNNTSDILATTIPFGIALGSLAFL 2380  
Qy 745 NEIKYPTKGDIIYDI-DTFSDTFGDCTPISINANI--NEQQSGKDTNTNGSETSDSPV 800  
Db 2381 FLKKKPKSPVLLRVLVDVHKDYGTPPKSSNRYIYPASDRHKGK----- 2425  
Qy 801 SHEPESDAAINVEKLSGDESSSETRGILDINDPDSVTNNVNEVHDASNTQGSVNTSDITN 860  
Db 2426 -----TYIYMEGDSGDEKYPMSV-----TTDVT 2451  
Qy 861 GHSE 864

```

Db      2452 SESE 2455

RESULT 28
ADR43499
ID      ADR43499 standard; protein; 3056 AA.
XX
AC      ADR43499;
XX
DT      04-NOV-2004 (first entry)
XX
DE      CSA-binding malarial variant surface antigen.
XX
KW      antimalarial; gene therapy; vaccine; Plasmodium falciparum; malaria;
KW      variant surface antigen; glycoaminoglycan; chondroitin sulphate A;
KW      immune response; erythrocyte; placental parasite; prophylaxis;
KW      pregnancy-associated malaria; diagnosis.
XX
OS      Plasmodium falciparum.
XX
PN      WO2004067559-A1.
XX
PD      12-AUG-2004.
XX
PF      30-DEC-2003; 2003WO-DK000938.
XX
PR      27-JAN-2003; 2003DK-00000102.
XX
PA      (UYKO-) UNIV KOENHANS.
XX
PI      Theander TG, Salanti A, Hviid L, Staalso T, Jensen ATR;
PI      Lavatsen T, Dahlbaeck M;
XX
XX      WPI; 2004-594166/57.
DR      N-PSDB; ADR43498.
XX
XX      New polypeptides and encoding nucleic acid molecules useful for
XX      diagnosing, preventing or treating pregnancy-associated malaria or for
XX      identifying agents capable of preventing or treating pregnancy-associated
XX      malaria.
XX
PS      Claim 1; SEQ ID NO 2; 138pp; English.
XX
XX      This sequence corresponds to a Plasmodium falciparum (malaria) variant
XX      surface antigen which binds to the glycoaminoglycan chondroitin sulphate
XX      A. The protein is designated vsa2csa. The protein is capable of inducing
XX      an immune response against a molecule expressed on the surface of an
XX      intact erythrocyte infected by a placental parasite. Sub-sequences
XX      comprise at least one B-cell epitope or one or more CAG-binding motifs
XX      and do not comprise a CIDR domain or DBL-gamma domain. The protein is
XX      gender specifically recognized or is recognized in a parity dependent
XX      manner. The polypeptide or nucleotide sequence is useful for
XX      manufacturing a composition that prophylactically or therapeutically
XX      reduces the incidence, prevalence or severity of pregnancy-associated
XX      malaria in a female subject. The composition and methods may also be used
XX      for diagnosing the above condition or for identifying agents capable of
XX      preventing or treating pregnancy-associated malaria.
XX
SQ      Sequence 3056 AA;

Query Match          9.8%; Score 637.5; DB 8; Length 3056;
Best Local Similarity 22.9%; Pred. No. 3.1e-30;
Matches 297; Conservative 165; Mismatches 462; Indels 371; Gaps 66;

QY      1 MKGYFN-IYFLPLIFLYNVRINESIIGRTLYNRQDESSDISRVNSP----- 47
Db      1854 MSYNYKII-----RKSNKESEDGKDYSMIMEPTVIDYLNKRCNG 1893
QY      48 ELNNNH-----KTNLYSDYEDVNNKLNSFVNKSVKKGKSLSPNNKTKSYDIIPPSY 102
Db      1894 BINGNYICCSCKNIGENSTSGTVNKLQKK-ETQCEDNKGPLDLMNKLNMVD---PKY 1948

103 SYRNDKFN--LSENEGNSGNTNSNFPANTSEISIG-KDNKOYTF---IQKRTHLFAC-- 154
1949 SEHKMKCTEYVLEHVE-----OLKEIDNAIKOYKLYPLDRCFDDKSKWKVCDL 1997
155 -----GKRKS-----IKW-----ICRESEKIT-VCVDPDKIOLCVANFLN--SRLETMEK 198
1998 IGDAGICKKHTKLDDELDEWNDVMDRDPYNKYKGLIPRRRLQCFSTRVIRGPANLRLKE 2057
199 FKEIFLISVNTKLLYN-KNEGKDPISFCNELNSESDFRSSFIGDDMDFGNTDRVKG 257
2058 FKEILKGAQSEGKPLGNYNEDKDEKALAMKNKSYDYIYIKGSDM----- 2106
258 YINTKFSDDYK--EKATVEK-ILNNIKK--EWEKKNKALNWNMIVNHK--GN--ISKCAI 308
2107 LTNIQFKDIKRLDLLEKETNNTKEVDWDETNNKKSIMWAMLCGYKSGNKKIDPSWCT 2166
309 IPABE--POINLWIKWENENFLMKRFLNLIKOC--VEN-KKYEACFGGRLUPCSSYT 363
2167 IPTTETPPQFLRWIKWGTNVCIOKEEHKEVYKSCSNVTNLGAQESSESKNCTSEIKKYQ 2226
364 SFMEKSKTOMEVLNLYKKNSGVD--KNFLNDLPKKNKNKNDLDDFPKNEKEYDDLDCD 421
2227 EWSRKRSIQWEAISSEGY-KKYKGMDEFKNTFNKIKEPDANEPNANEYL---KKHCKRCP 2282
422 -----RYTATIIKSFNGPAKNDVDIASQIN-----VNDLRGF 454
2283 GFNDMQEITKYTNIGNEAFKQ--IKEQVDIPAELEDVIYRLKHHEYDKGNDYICNKYKI 2340
455 GCNYKSNNEKSN--CTGTFTNKPFGTCEPPRRQTLCLGRTYLLHRGHEBD----- 503
2341 NVNMKNQNDTWTDLVKNSSDINK--GVLLPFRKFL-----KIDESDICKYKRD 2391
504 --YKEHLLGASIVEAQLLYKYEKENALCSIIQNSYADLADIIKGSDIK-----DYVG 557
2392 KLFKDFIYSSAISSEVERLKKVYGEA-KTKVVMHAKYIFADIGSIKGDMMENNSDDKIG 2450
558 KMEENLNKVNKKRNEESLKIIFREKWDNENKENVKMSAVLNK--ETCKDYDFKQK 615
2451 KILGDVG-----QNEK-----RKWWDNWKYHIMESMLCGYKHAYGNISENDRKMD 2498
616 IP-----QFLRWFKWGGDDPCFKRKGKIYSFESFKVECKKOC-----DENTCKNKC 662
2499 IPNNDDEHQFLRWFPQEWTFNCTKRNE---LYENNVTACNSAKCNTNSGSDVCKECTEAC 2555
663 SEYKWKIDLKSEVEKQVDKYTKDNKKMYDNIDEVKNKEANVYLKESK---ECKDVNF 719
2556 KNYSNFILIKKKYQSLNSQY--DMNYK-----TKAEKKSPFYFKDKNGECSCLSYF 2609
720 DDKIFNESPNEYD-----MCK-----KCDIYKLYNEIK 748
2610 XDETRWKNPYETLDDTEVKNNCKCKPPPPASNNNTSDILOKTIIPGIALALGSIAPFMK 2669
749 YPKTKHDIYDI-DTFSDTFGDTPIISINANI---NEQOSGK-----DTSNTGNS---- 793
2670 KPKTPVLLRLVDIPKDGIGIPTPKSSNRYIPYASDRYKGYKIYMEGDTSGDDDKYWD 2729
794 -ETSDSPVSHPEPDAIN-----VEKLSGD 818
2730 LSSSDITSSSESEYEEVDINDIYVPSFPKYKTFIBLVLEPSKRDFTNTSSGDTFTNKLTD 2789
819 ESSSETRGILDINDPSTVNNVNE---VHDASNTQGSVNTSD---ITNGHSE----- 864
2790 EWNQOLKODFIE---QYLNQIKQDFILHDSWDEKFPITQDRFLDSHEEVTYNDWNV 2845
865 -SSLNRTTNAQD-----IKIGRSGNEQSDNOENSSHSDNSGSIITIGQVP 908
2846 PENINRITNNMDDPKYCSNNMYTGTDLINDSLNGQYDIYDEMLKRENE---LFGTYH 2902
909 SEDNTQNTYDSNP-----HRTFPNALASLPDSD---KLINEIEGPDSS 948
2903 TKYTFNFSVKQTPSDPIINQLDLYHKWIDKHRDICEQWKT--KEDMLYKLNBYNMWERK 2960
949 RDSENGRGDTTSTNTHDV--RRTNIVSERRVNSH-----DFI-----RNG 985

```

```
Db 2961 EYLLDIOPSLDDIHKINDEYNIISTNNIYDHPQSQTPIQLLGSTNIISYITTEQNG 3020
Qy 986 MANNNAHQYITQENNGIRGQ---BESAGNSVN 1017
Db 3021 LRTNISMDTYIDETNNNNVATSIIGDDQMSNVN 3055

RESULT 29
AAB62142
AC AAB62142 standard; protein; 3542 AA.
AC AAB62142;
XX
XX 29-MAY-2001 (first entry)
Dt
XX
XX P. falciparum FCR3.varCSA protein.
XX
XX FCR3.varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1;
KW erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
KW malaria; protozoacide.
XX
XX Plasmodium falciparum.
XX
XX WO200116326-A2.
XX
XX 08-MAR-2001.
XX
XX 01-SEP-2000; 2000WO-US024195.
XX
XX 01-SEP-1999; 99US-0152023P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
PI Gysin J, Pouvelle B, Fujii N, Smith J;
XX
XX WPI; 2001-235109/24.
XX
XX N-PSDB; AAF57301.
XX
XX Novel FCR3.varCSA protein, useful for modulating parasitized red blood
XX cell binding, sequestration and onset of maternal malaria.
XX
XX Claim 12; Page 63-71; 78pp; English.
XX
XX The invention relates to a P. falciparum FCR3.varCSA protein, that is
XX capable of binding to chondroitin sulfate A (CSA). The var gene and the
XX corresponding P.falciparum erythrocyte membrane protein 1 (PFEMP1)
XX modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
XX protein and the encoding gene are useful for treating and preventing
XX maternal malaria in a patient afflicted with maternal malaria. The
XX present sequence represents the P. falciparum FCR3.varCSA protein
XX
XX Sequence 3542 AA;

Query Match 9.3%; Score 601.5; DB 4; Length 3542;
Best Local Similarity 20.8%; Pred. No. 6.9e-28;
Matches 303; Conservative 192; Mismatches 466; Indels 499; Gaps 68;

Qy 56 NIYDSYEDVNNK---LINSFVENKSVK---KRSLSFI-----NNKTKSYD- 96
Db 1118 NIISDKYKELHQMSVNSGIISSASTAKNHIDRNVIPLSELYQQGKSNKSGTSDE 1177
Qy 97 --IIPPSYSYRNDFKNSLSENEEDNSGN---TNSNNFANTSEISIGKDNKYTFIQK-RT 149
Db 1178 SAVIGTNTYEN-----VGAYLHDTGTFDDCQSQNEFCDD--EKSDGDKNEKYAFRDKPD 1230
Qy 150 HLFACGKIR-----KSTKWTICRENS----- 169
Db 1231 HDGACGCKSGSKPTRVQIKTKKKAEBKDTECKTVNDILKENDGKKQVEDCHPKKNSNGYP 1290
Qy 170 -----EKITVCVPDRKIQLCVANFLN---SRLETMEKFEKFLISVNTAE--KL 213
```

```
Db 1291 DMOCGNINLVEDPRVCMPPRRQKLCVHFLEANDNEIKKLSQVNLKEAFISAAAEFFSW 1350
Qy 214 LYNKN-----EGKOPSIFCNELRNSFSDFRSFISGDDMDFG-GNTDRVKGYI 259
Db 1351 YYYKSKDGBEGNLDKELKEGKIPPAFLRSNFFYFGDYRDLFTGDISKGHGESKLEQEI 1410
Qy 260 NTKFSDYYKEKNVEKLN---NIKKEWKNKANLWNHMIVNHKGNISKECAIIP----- 310
Db 1411 DSLF-----KNGDQKSPNGKTRQEWTEHSHEWEAML-----CALVKIGAKKD 1454
Qy 311 -----AEEPJNLWLKEWNEFLMEKKRFLNLNIKQCVENKK 347
Db 1455 DFTENGYNNVFSKSTLEBFAKPOFLRWITWYDDYCYTRQKYLKDVQSKCKSNDQ 1514
Qy 348 YEACFGGCRLLPCSSYTSFMKSKSTOMEVLTNLYKKNSGVDKNNFLNDLFFKNNKNDLDD 407
Db 1515 LK-CDTECNKKCEDYVYMKKKKEWIP-QDKYKDER---DKKRF----- 1554
Qy 408 FFKEKEYDDLCDRYTATIISFLN-----GPAKNVDVDIASQVNVNDL----- 451
Db 1555 ----DRQHIGVMVTDYTGNTAIDYLNKFTASCGDKPGSASVV-QRNQLLEKQAYYDAD 1609
Qy 452 RGFGC-NYKSNNKESWN-----CTG-----TFTNKFPGTCE-----PP 483
Db 1610 KHGCTKFIENDDKYTNISSKOKCKGLVKEANTGAIKWQNKGNPNYNNLKELTEDVLFPS 1669
Qy 484 RRQTLCLGRTYLLHRGH-----BEDYKEHLGASIEAQLLKYYKEKEDNAL----- 531
Db 1670 RRLRICF---HALDGNVTDPEVKDENGRLKRLMEVAATEGYNLGOYYKEKEKIKTSD 1726
Qy 532 -----CSIIQNSYADLADIIGSDIHKDYGGKMEENLNKV-NKD-----KRNNEE 576
Db 1727 AHKYSYEVPPCSAMKYSFYDLRDIILGIDNLED-EKQKTEENLKIFNKNGTSGVGKSDS 1785
Qy 577 SL----KIREKWNDEKNVWKMVASVLK-----NKETCKDYDFQKIP----- 617
Db 1786 TTGNPGSTARKFPWNEKNKBCVWNAMICGYKRGDRDNGSGNSARSDEDLKKGSGVSDDDY 1845
Qy 618 -----QFLRWFKEWGDDFCBKEKEKIYSPESPKVECKKDCDEN-----TKKN 660
Db 1846 PMGKNRDEGTAYQFLRWFAEWGEDFC-KHKEK--ELEKLVGACNDYTCGNEDKPKKCTD 1902
Qy 661 KCSEYKWIIDLKSEYEKQVDKYTKDNKKMYDNIDEVKNKB--ANYLKEKSKCEKDVNF 719
Db 1903 ACTQYKKFISEWKPOVEKQIKKYGENKD-KIYSEHPVAKDAEDAREYLDKQLK----- 1954
Qy 720 DDKIFNESPNEYEDMCKCDEIKYLNKIKYPKTKHDIYDIDTFTDFTGDCGTPISINANIN 779
Db 1955 --KICENKSGDCBYKCMK-----DVSTQRLTDGNSQNNMPASLD 1990
Qy 780 EOQSGKDTNTGNSETSDSP--VSHEPESDAAINVEKLSGDESSSETRGILDINDP--SVT 836
Db 1991 DEP--KEVEGKNCQVPRGPPVRVRETPTSPRVSLSKATA--SKKEAKTAPPTKQPKVE 2046
Qy 837 NNVNEVHDASNTQGSVNTSDITNGHSESSLNRTTNAQDIKIGRSNGEQSDNQENSSHSS 896
Db 2047 NLTTENRAQTRTRAAQQRKRTS-----TATTTESD--VGTMVKAILLSNKPDSRGGI 2097
Qy 897 DNSGSLTIGQVP-----SEDN-----TONTYDS----- 919
Db 2098 EGCNPKTYGQYPKWGCIVGKSEKENNGICMPPRRKKLCINNIQYLVNTEYENKKNNDIKEA 2157
Qy 920 -----QNPHRDTFNALASLPD-----DKINEIEGFSDSRDS 951
Db 2158 FIKCAAIEQTFLWKVLIENPAEANELQGTIDPEFKRIMYYTYGDKDFGFGTIDISNDK 2217
Qy 952 ENGRGDTTNTHDVRTNIVSERRVNSHD-----FIRNGMANNNAHQ 994
Db 2218 K-----IITVTSV--TTILNENNNKKQDKKDEELRKIPWEKNKKFIFWEGMYGLTYH- 2269
Qy 995 YITQIENNGIIRQOESAGNSVNYKD----NPKRSNFSSENDHKNIQYNSRDTKRVRE 1050
```

Db 2270 -LTD-EN-----EKEKIRDNYQNDMTKLTPLSLSEFVKRQFLRWFTWEAEFCNKKE 2321  
 Qy 1051 BIIKLS---KQKCNNEYSMEYCTYSDBRNSPGPCSRERKCLCCQISDYCLKYFNFS 1107  
 Db 2322 QLLKLEAGCKEYECNG-----SNDGKTQEC---AEACVYQNF-- 2356  
 Qy 1108 IBYNYCIKSEIKSPYKCFKSEGQSSIPYFAGGILVIVILLSSASRMGKSNBEYDIGE 1167  
 Db 2357 -----IKKWKTEYERQREKF-----KKDKGKKYKDYDPSTE 2387  
 Qy 1168 SNIETATPENNLYN-KLSRI 1186  
 Db 2388 RDIEKATCAHEYLNNKLKEL 2407  
 RESULT 30  
 AEB22180  
 ID AEB22180 standard; protein; 3147 AA.  
 XX  
 AC AEB22180;  
 XX  
 DT 22-SEP-2005 (first entry)  
 XX  
 DE Codon optimized P. falciparum VAR4.  
 XX  
 KW antimalarial; vaccine; protein purification; pharmaceutical; diagnostic;  
 KW plasmodium falciparum infection; infection; VAR4.  
 XX  
 OS Plasmodium falciparum.  
 XX  
 FN WO2005063804-A1.  
 XX  
 PD 14-JUL-2005.  
 XX  
 PF 28-DEC-2004; 2004WO-DK000922.  
 XX  
 PR 30-DEC-2003; 2003DK-00001954.  
 XX  
 PA (UYKO-) UNIV KOBENHAVNS PANUM.  
 XX  
 PI Jensen ATR, Hviid L, Jorgensen L, Lavsten T, Magistrado P;  
 PI Nielsen MA, Salanti A, Staalso T, Theander TG;  
 XX  
 DR WPI; 2005-522559/53.  
 DR N-PSDB; AEB22179.  
 XX  
 PT New isolated VAR4, VAR5, and VAR6 polypeptides isolated from Plasmodium  
 PT falciparum, useful as vaccines for treating or preventing severe malaria.  
 XX  
 PS Disclosure; SEQ ID NO 8; 193pp; English.  
 XX  
 CC The invention describes an isolated polypeptide comprising an amino acid  
 CC sequence selected from at least one of SEQ ID NOS 2, 4, and 6 (3553, 2994  
 CC and 3346 amino acids) for use as a medicament. The polypeptides, nucleic  
 CC acids and vectors are useful as vaccines for preventing and/or treating  
 CC malaria. The malaria is preferably severe malaria caused by Plasmodium  
 CC falciparum. An in vitro diagnostic method is useful for detecting anti-  
 CC malarial antibodies. The polypeptides are useful for generating a vaccine  
 CC against severe malaria, where the vaccine comprises antibodies generated  
 CC against the polypeptide which have cross-reactivity to parasites causing  
 CC severe malaria and the ability to inhibit adhesion to endothelial cells.  
 CC The polypeptides are also useful in screening for inhibitor molecules  
 CC capable of inhibiting binding of any of the polypeptides to a receptor  
 CC expressed on endothelial cells. This is the amino acid sequence of  
 CC Plasmodium falciparum Var4 that is transcriptionally upregulated in  
 CC Plasmodium following increased antibody recognition and is encoded by a  
 CC codon-optimized polynucleotide.  
 XX  
 SQ Sequence 3147 AA;

Query Match 9.1%; Score 592; DB 9; Length 3147;  
 Best Local Similarity 20.0%; Pred. NO. 2.3e-27;  
 Matches 315; Conservative 205; Mismatches 511; Indels 544; Gaps 68;

Qy 39 SDISRVNPELNNHNTNIYDSYEDVNNKLINSFVENKSVKKRSLSFNNKTKSYDII 98  
 Db 1252 TDVAKILQGEAN-----ETMLKNSNGNDKDSKLGKAEED-- 1289  
 Qy 99 PPSYSYRNDKPNLSUENEDNSGNTNSNFANTSISIGKONKQYFTQKRTHLFACGIKR 158  
 Db 1290 ---YS-RGGTSPDFNNNLGIGTQKSHANDSQQPCYCKQQRNFV----- 1331  
 Qy 159 KSIKWI CRENSEKIT--VCVPDRKIOLCVANFLNSLETM--EKFKEIF----- 203  
 Db 1332 -GTEWSPKDNHRKTHPEAYMPPREHICTSN-----LEYLIHRKKPIIEGDPNKIHS 1385  
 Qy 204 -----LISVNTA---KLLYNKNEG-KDPSIFCNELRNSPDRSSFIGDDMDGEGNTDR 254  
 Db 1386 LLGDVLLAAKYEAEENIKLYEENNRKQDEGICRAMKYSFADIGDIIRGKDM-WIENDA 1444  
 Qy 255 VKGIYN-----TKFSD-----YYKEKQVKEKLNKKEWKEKANLWNHMIWNHG-NIS 303  
 Db 1445 KRLQTNLKEIPTKIKERTGGTTYNEDNDPYL-KLRADWWEANRAKVKAMCKCTNGVDIT 1503  
 Qy 304 KECALIPAEB--POINLIWIKENENFLMEKKRLFLNIKDCVSE---NKKYEACF----- 352  
 Db 1504 CDSHTPLDDYIPORLWMTWEAWCYKASQETKLEEKCSQKSGKSGNECYRETYE 1563  
 Qy 353 -GGCRLPCSSYTSFMKSKSKTQMEVLTN---LYKKKNSGV-----DKN--NFLN 394  
 Db 1564 CNDCKQACEEYKRIKTWADQWKVSNKYEDLYKKAQNPNAVLKDNKDEKDKVDFLT 1623  
 Qy 395 DLFKKNN-----KNDLDD-----KNDLDD---FFKNEKE-Y 415  
 Db 1624 QLKANNGEKTGVHTVYSTAAGYIHOEARTRECOBREFCDKKGIDNTSYAFKDPHGY 1683  
 Qy 416 DDLCDCRYTATIKSFLNGPAKNDVDIASQINV-----NDLRGGF---CNVK-----S 460  
 Db 1684 ATACDC-----INRSQTEEPKKEENVESACKIVVEVLSRPRDKTTGIDHCNPKYPRK 1738  
 Qy 461 NNEKSNWCT-GTFTNKFPFGCEPPRRQTLG-LGRTYLLHRGHEEDYKHEHLGASIEAQL 518  
 Db 1739 ENYFGWNTCPQFSKHAGACMPRRRIKLCVINQYLNKESPEELKAFQCAAIETW 1798  
 Qy 519 LKYYKKEKDNALGSIIONS-----YADLADIIGSDIHKDYGK---KM 560  
 Db 1799 LMQYKKDKNGVAQAALNSGTIPDDPKRQMYFTFGDYRDLCLDTDISSKADTSTGVGKV 1858  
 Qy 561 EENLKV-NDKKNEESLKIIFREKWDENKENVKWSAVLKNKEICK--DYDKFKKI- 616  
 Db 1859 KINIDSVFQKIDITNVEQ-----RKPWNGKNAEATWDGMLCALSYNTTKNNMDYNAHTKLN 1914  
 Qy 617 -----POFLRPFKEWGDGDFCEKREKIYSFESFKVECK----- 649  
 Db 1915 PTYGYNAIKSELEDFVNRPOFLRWFTWSEDFCTERSIKIKELETKCNDCTVSESGTSDA 1974  
 Qy 650 --KKDCDE---NTCKNKSEYKQWIDLKSEYKQVDYTKDKNKKMYDNIDV-RNKE 702  
 Db 1975 TGNKTDCKDKCDBCKRACCTYKTLKWKMTQYKQSKYFDDKRKELYSIDDVASSTQ 2034  
 Qy 703 ANVYLKESK-----FCKDVNFDKIFNESP-----NEYEDMCKEC----- 738  
 Db 2035 AYQYLHAQLKKLQGNADCKCMGESKETQGPNSHSHMPASLDDEPEVNGKCNCKVK 2094  
 Qy 739 -----DEIKYLN-----EIKYPKTKHD 755  
 Db 2095 HRPOPPLALPPAPSGPPASDQIEHNRGSRSGDQGLPARPPPPQAAPPQPKPKRT 2154  
 Qy 756 IYDIDTFSDTFGDG-----TPISINANINBQQSGKTSNTGNSFSDSPVSHPESDAAI 810  
 Db 2155 -----GEGJGRNLPPADRNLTNLSDEEEDDDDDDEVQSEETPPSEAEGEHV 2203  
 Qy 811 NVEK-----LSGDES----- 820  
 Db 2204 ETEBETPKVKEKTEGAGATEVTKQGSAPTATPTTPTVEDICATVAKALGDKSLNAACALKY 2263

QY 821 ----- 820  
 Db 2264 GKNNRLGWKIPSGDKTDSNGAPRRARSAGGKSDSEKSGICVPPRRRLYIKKIV 2323  
 QY 821 ---SSETRGILDIN-DPSVTNNVNEHDASNTQGSVNT-----SDITNGHSESS----- 866  
 Db 2324 DWAESQKTVTSVNGDNGQSVVNGASGSGSGTQSDASDVSGNGASTSPOVAL 2383  
 QY 867 LNRRTNAQDI-----KIGRSGNQSDNQENSSH-----SSDNGSLTIGQVPSSED 911  
 Db 2384 LHAPVKSAAIETFPFAWHKYVDKE-IEEKEQAAQHLVQRKTSNPQKLEGEIP-ED 2441  
 QY 912 NTQNTYDSQPHRD-----TPNALASLPDDKINEI-EGFDSRDSEN-----GRGDTTS 960  
 Db 2442 FXRQMFYTLGDYRDLVGDKTMIETIALEKSGDGTIEDISEKIPKILDGNNKAAGGPKQP 2501  
 QY 961 NTHDVRRTNIVSERRVNSHDFIRNGM-----ANNNAHQYITOI-----ENNGIIRQEE 1010  
 Db 2502 NSGKTPQ-----EWWKENAKHIIWGMICALTYNTDSNGKDKKIQQVKATONTDLFQKLKK 2556  
 QY 1011 -----SAGNSVNYKDNPKRSNFSSENHKKNIQOYNSRDTKRVRBEEIIKLSKQNK 1060  
 Db 2557 DNDYETVSFGASGTGAKSNDOTKLKNFVVRPTYFRWLEEWEEBFCRQKHLYIHKDCR 2616  
 QY 1061 CNNEYSMEYCTYSDERNSSPGCSREERKKL-CCQISDYCLKYFNFYISYIYNCIKSEIK 1119  
 Db 2617 DN-----KFCSGDGLRCDKVPDKKIDFKHFDPCSCARHCRSYRKWI-----ERK 2661  
 QY 1120 SPEYKCFKSGQSSIPYFAAGILVIVILLSSASRMGKN-----EYDYGESN----- 1169  
 Db 2662 KTEVE--KQZ-----SAYSQKSNVYNGSGDGNNDKEFYT 2697  
 QY 1170 -IEATFEENNYLNKL 1183  
 Db 2698 KLETCTKATNFLESJ 2712  
 RESULT 31  
 AEB22174  
 ID AEB22174 standard; protein; 3553 AA.  
 AC AEB22174;  
 XX  
 XX 22-SEP-2005 (first entry)  
 XX  
 XX Plasmodium falciparum VAR4.  
 XX  
 XX antimalarial; vaccine; protein purification; pharmaceutical; diagnostic;  
 KW Plasmodium falciparum infection; infection; VAR4.  
 XX  
 XX Plasmodium falciparum.  
 OS  
 XX WO2005063804-A1.  
 XX  
 XX 14-JUL-2005.  
 PD  
 XX 28-DEC-2004; 2004WO-DK000922.  
 XX  
 XX 30-DEC-2003; 2003DK-00001954.  
 XX  
 XX (UYKO-) UNIV KOBENHAVNS PANUM.  
 PA  
 XX Jensen ATR, Hviid L, Jorgensen L, Lavsten T, Magistrado P;  
 PI Nielsen MA, Salanti A, Staalso T, Theander TG;  
 XX  
 XX WPI: 2005-522559/53.  
 DR N-PSDB; AEB22173.  
 DR  
 XX New isolated VAR4, VAR5, and VAR6 polypeptides isolated from Plasmodium  
 PT falciparum, useful as vaccines for treating or preventing severe malaria.  
 XX  
 XX Claim 1; SEQ ID NO 2; 193pp; English.  
 PS  
 XX

CC The invention describes an isolated polypeptide comprising an amino acid  
 CC sequence selected from at least one of SEQ ID Nos 2, 4, and 6 (3553, 2994  
 CC and 3346 amino acids) for use as a medicament. The polypeptides, nucleic  
 CC acids and vectors are useful as vaccines for preventing and/or treating  
 CC malaria. The malaria is preferably severe malaria caused by Plasmodium  
 CC falciparum. An in vitro diagnostic method is useful for detecting anti-  
 CC malarial antibodies. The polypeptides are useful for generating a vaccine  
 CC against severe malaria, where the vaccine comprises antibodies generated  
 CC against the polypeptide which have cross-reactivity to parasites causing  
 CC severe malaria and the ability to inhibit adhesion to endothelial cells.  
 CC The polypeptides are also useful in screening for inhibitor molecules  
 CC capable of inhibiting binding of any of the polypeptides to a receptor  
 CC expressed on endothelial cells. This is the amino acid sequence of  
 CC Plasmodium falciparum Var4 that is transcriptionally upregulated in  
 CC Plasmodium following increased antibody recognition.  
 XX  
 SQ Sequence 3553 AA;  
 Query Match 9.1%; Score 592; DB 9; Length 3553;  
 Best Local Similarity 20.0%; Pred. No. 2.8e-27;  
 Matches 315; Conservative 205; Mismatches 511; Indels 544; Gaps 68;  
 QY 39 SDISRVNSBELNNHKTNIYDSYEDVNNKLINSFVENKSVKKRSLSPINNKTYSVDII 98  
 Db 1252 TDVAKILOGEAN-----ETMLKNSNGNDKDESKLGAEGD-- 1289  
 QY 99 PPSYRNDKFNLSLSENDNSGNTNSNFANTSEISIGKDNKQYTFIQKRTLHFACGIKR 158  
 Db 1290 ---YS-RGGTSPDFNNLGGITQKSHANDSOQPCYCKDKQFNV----- 1331  
 QY 159 KSIKWICRENSKIT---VCVPDRKTQLCVANFLNSRLTM-EKPEKIP- 203  
 Db 1332 -GTSEFKDNHRKTRTHPEAYMPPRREHICTSN-----LEYLIHKRKKPIEGDPNKKIHS 1385  
 QY 204 -----LISVNTAE---KLNYKNKG-KDPSIFCNELRNSPSPRSFIDGDMDFGGWTD 254  
 Db 1386 LIGDVLAAKYBAENIKLYEENNNRKKQDEGICRAMKYSFADIGDIIRGKDM-WIENDDA 1444  
 QY 255 VKGYIN-----TKFSD-----YYKEKNVKEKLANIKKEWKEKNANLWNHNVNKG-NIS 303  
 Db 1445 KRLQTNLKEIFYKIEKTGTYINEDNDPYL-KLRADWEANRAKWKAMCKKTNGVDIT 1503  
 QY 304 KECALIPABE--PQINLWIKENENFLMEKKRFLMKKCVB-----NKKYBACF----- 352  
 Db 1504 CDSHTPLDDYIPQRLRWMTAEAWYCKAQSOEYKLEEKCSQCKSGKGGNECYRETKE 1563  
 QY 353 -GGCLPSSSYTSFWMKSKTOMEVLTN---LYKKNSGV-----DGN--NFLN 394  
 Db 1564 CNDCKQACBEYKRIKTWADQWKVISNKYEDLYKGAQNPNTNAVLKDKNKEKKNVIDFLT 1623  
 QY 395 DLFKKNK-----KNLDD-----FPKNEKE-Y 415  
 Db 1624 QLOKANNGEKTVHTVYSTAAGYIHOEARTRECQOEFECDKNGKDINTSYAKDPHGY 1683  
 QY 416 DDLCDRCRYTATIKSPFNGPAKNDVDIASQINV-----NDLRFG-----CNYK-----S 460  
 Db 1684 ATACDC-----INRSQTEEPKKEENVESACKIVEEVLSPKPRDKTTGGIDHCNPKYPRK 1738  
 QY 461 NNEKSNKCT-GTFTNKFPCTCEPPRQTLCLGRTYLLHRGHEEDYEKHLGLASIVEAQL 518  
 Db 1739 ENYFGWNCPTFGQPKSGHAGACMPPRRIKLCVINQLYNLEKKSPEELRKAFIGCAAIETYW 1798  
 QY 519 LKYKYEKEDENALCSTIONS-----YADLADIIGSDIIXDYDGK-----KM 560  
 Db 1799 LAQKYKKNXGGVAQAKLNSGTIPDDFKQMFYTFGDRDLCLDTDISSKADTSTGVGVK 1858  
 QY 561 EENLNKV-NKDKRNEESLKIIFREKWDENKENVKNSAVLKNKETCK--DYDKFOKI- 616  
 Db 1859 KINIDSVFQKIDITNVEQ---RKPWWGKNAEAWDMLCALSYNTTNKMDYNAHTKLN 1914  
 QY 617 -----PQFLRWFKEWGDDPFCEKKEKTIYSPESPKVECK----- 649  
 Db 1915 PTYGVNAIKSELEDVFNRFQFLWFTWSEDFCTERSIKIKELETCKNDCVTSESGETSDA 1974

QY 650 --KKDCDE---NTCKNKCSEYKWIIDLKSEYKQVDKYTKDKNKKWYDNIQV-KNKE 702  
 Db 1975 TGNKTCDDKCKDECKRACCTYYKWLKNWKYQYTKQSKYFDDKXKELYKSIDVASTQ 2034  
 QY 703 ANVILKEKSK-----ECKQVNFDDKIFNEBP-----NEYEDMCKKC----- 738  
 Db 2035 AYQVLAQLKLCGNADCKMDGSKETTTGQPDNSHSHMPASLDDEPEEVNGKCNCKVK 2094  
 QY 739 -----DEIKYLN-----EIKYKTKGHD 755  
 Db 2095 HROPPLALPPAPSGPPAEQIEHNRGRSERGGQPLPARPPPPQAAQPPQPKPRT 2154  
 QY 756 IYDIDTFDTCGDG-----TPISINANINEQQSGKDTNTGNSSETSDSPVSHHEPESAAI 810  
 Db 2155 -----GEGLRNLPADRNTNLSDEEEDDEDDVQEEETPPSEAGEGHV 2203  
 QY 811 NVEK-----LSGDES----- 820  
 Db 2204 ETEBETPKVKEKTEGAGATEVTKQGSAPTATTPTTVEDICATVAKALGDKSLNAACALKY 2263  
 QY 821 ----- 820  
 Db 2264 GKNSRLGWKCIPTSGDKTDTSENGAPRRASAHGKSDSEKSGICVPPRRRLYIKIV 2323  
 QY 821 ---SSETRGILDIN-DPSVTNNVNEVHDASNTQGSVNT---SDITNGHSESS----- 866  
 Db 2324 DWASQSKVTIVSNGDNGSDEVVNGASGSGSGTESQADSVQSGCASTSPQVAL 2383  
 QY 867 LNRTNAQDI-----KIGRSGNEQSDNQENSH-----SDNSGSLTIQGVPSD 911  
 Db 2384 LHAFKSAAIETFFAWHKYKVDKE-IEEKEQAQNLVORKTSENPOKLEGGEIP-ED 2441  
 QY 912 NTQTYDSQPHRD-----TPNALASLPDDKINEI-EGFDSRDSN-----GRGDTTS 960  
 Db 2442 FKQMFYTLGDRDILVGDKTMI EALEKSGDTKIEDISEKIPKLDGNNKAAGGPKQP 2501  
 QY 961 NTHDVRTNIVSERVNSHDFIRGM-----ANNAHVOYITQI---ENNGIIRGOE 1010  
 Db 2502 NSGKTQ-----EWWKENAKHIWGMICALTYNTDSNGKDKKIQOVKATDNTDLFQKLK 2556  
 QY 1011 -----SAGNSVNYKDNPKRNSFSENHDKNQIYNSRDKRVREBIKLKQNK 1060  
 Db 2557 DNDYETVSFGASGTGAKSNDTDLKNFVVRPTFYRWLEWGEFEFCRQKHKLYIKKDCR 2616  
 QY 1061 CNEYSMEYCTYSBERNSSPGCSREERKLI-CQOISDYCLKYFNFTYIEYNYCKSEIK 1119  
 Db 2617 DN-----KFCSGDGLRCDCEKVPDKDIKPHFDCPSCARHCRSYRKWI-----ERK 2661  
 QY 1120 SPEYKCFKSEGSSTIPYFAGGILVIVLLLSASRMGKSN-----EEYDIGESN----- 1169  
 Db 2662 KTEVB--KQE-----SAYSKOKSVNGSGDGGNNNDKEFYT 2697  
 QY 1170 -IEATFEENNYLNLK 1183  
 Db 2698 KLETCTKATNFLESL 2712

## RESULT 32

AE22176  
 ID AEB22176 standard; protein; 2994 AA.  
 XX AC AEB22176;  
 XX DT 22-SEP-2005 (first entry)  
 XX DE Plasmodium falciparum VAR5.  
 XX KW antimalarial; vaccine; protein purification; pharmaceutical; diagnostic;  
 XX KW plasmodium falciparum infection; infection; VAR5.  
 XX OS Plasmodium falciparum.  
 XX

FN WO2005063804-A1.  
 XX 14-JUL-2005.  
 XX 28-DEC-2004; 2004WO-DK000922.  
 XX 30-DEC-2003; 2003DK-00001954.  
 XX (UYK-) UNIV KOBENHAVNS PANUM.  
 XX Jensen ATR, Hviid L, Jorgensen L, Lavsten T, Magistrado P;  
 XX Nielsen MA, Salanti A, Staalso T, Theander TG;  
 XX WPI; 2005-522559/53.  
 XX DR N-PSDB; AEB22175.  
 XX PT New isolated VAR4, VAR5, and VAR6 polypeptides isolated from Plasmodium  
 XX falciparum, useful as vaccines for treating or preventing severe malaria.  
 XX PS Claim 1; SEQ ID NO 4; 193pp; English.  
 XX CC The invention describes an isolated polypeptide comprising an amino acid  
 CC sequence selected from at least one of SEQ ID NOS 2, 4, and 6 (3553, 2994  
 CC and 3346 amino acids) for use as a medicament. The polypeptides, nucleic  
 CC acids and vectors are useful as vaccines for preventing and/or treating  
 CC malaria. The malaria is preferably severe malaria caused by Plasmodium  
 CC falciparum. An in vitro diagnostic method is useful for detecting anti-  
 CC malarial antibodies. The polypeptides are useful for generating a vaccine  
 CC against severe malaria, where the vaccine comprises antibodies generated  
 CC against the polypeptide which have cross-reactivity to parasites causing  
 CC severe malaria and the ability to inhibit adhesion to endothelial cells.  
 CC The polypeptides are also useful in screening for inhibitor molecules  
 CC capable of inhibiting binding of any of the polypeptides to a receptor  
 CC expressed on endothelial cells. This is the amino acid sequence of  
 CC Plasmodium falciparum Var5 that is transcriptionally upregulated in  
 CC Plasmodium following increased antibody recognition.  
 XX SQ Sequence 2994 AA;

Query Match 8.6%; Score 554.5; DB 9; Length 2994;  
 Best Local Similarity 21.4%; Pred. No. 4.9e-25;  
 Matches 286; Conservative 172; Mismatches 425; Indels 451; Gaps 65;  
 QY 114 ENEDNSGNTSNFANTSEISIGKDNKQYFIQKTHLFA---CGIKRSIKWICRENSE 170  
 Db 839 KEEDTDKQVPPAATTPGVKPCDIVEKFKKHDNTGALDHCPKQDYPW---KNDK 895  
 QY 171 KIT-----VCVPDRKIQLCVANFLNSRLTMEKFEFLISVNTBAKLLYNK----- 217  
 Db 896 SLVDEGVTMPPRQKLCVINLEHFKEKTSDDLREAFIKCAAETILLWQYKEDNNGE 955  
 QY 218 -----NEGKPSIFCNELNSPSDFRSSFIDGMD-FGGNTDRVKGYINTKFSDYKEK 270  
 Db 956 DLQNLKSGKIPEDFKRQMFYTFGDRDPLFGTDISKLNKHTKHTAVKTNIDRIFPP----- 1010  
 QY 271 NVEKLNN-IKKEWKKNAKLNWNHMI--VNHKGNISKECALIP----- 310  
 Db 1011 -TERTNDTIRKEFWERKNAESIWOGLMCALSTNSKDKDDVDQKELNSTNYDTIKNLE 1069  
 QY 311 --AEPEQINLWIKENENFLMEKKRFLNIDKCVENKCYEACF-----GGCRLPCSS 361  
 Db 1070 DFANRPQFLRWFIEWSDFCERKKEKVGSA--KNDYEGCANTKDNNGNCVACNA 1127  
 QY 362 YTSFMKSKTKQMEVLTNLY---KKQN-----SGVDKNNFLND-----LFFKN 401  
 Db 1128 YKTYITDKGEYKQAKKFDIDKSNKPGYEDYSGKASEYLKKEKINSSCDYMLKLDN 1187  
 QY 402 KNDLDDFFKKEKEDD-----LCDCRYTATII--KSFINGPAKNDVDIASQINVDLRGF 454  
 Db 1188 SNYWE---KPHTTYDDNSLQNKSCPLSPCEIVDKTLGDKTSKSYAE----- 1231  
 QY 455 GCNYKSNNEK---SWNCTGT--FTNKPFGTCEPPRRQTLCLIG--RTYLH----- 497

Db 1232 GCKWKYKMPGLGLWCNDKEGKEDGLCLIPRRKRLVYKDLTSPDHTTVCGLREAFI 1291  
Qy 498 -----RGHEBDEYKHL-----LGASIIYEAQLLYKYKEDENALCSIIQN 537  
Db 1292 KCAAVETFAWHEFTKEKREYKEKQNGELGFIENDQIPK-----DPDNPNQKIRKN 1346  
Qy 538 -----SYADLADIIGSDI-IDYGYKMEENLNKYNKDKK-RNEESLKIFR 582  
Db 1347 GEIHEEFKSQMFYTLADILFLGNIGIGNDMG-KVKSNDIKVFANSSGKTPTAKKTP 1405  
Qy 583 EKWDENKENVWVMSAVL-----KNKETCK-----DND 611  
Db 1406 KEWEKNAKOIWEGLCALSYDTTKIKNEELRKLIDPKNSVNYVEKVTFFSDNNTNLS 1465  
Qy 612 KFKIPIOLRPFKPGWDDCEKREKIKYSPESFVCK-----KKOCD-----E 655  
Db 1466 KFTERRPPFRWFQWGEFCRKKIKIDKIEK-----ECRGPVGRNHCDGDFDCSEIGPNE 1522  
Qy 656 N-----TKNKKSEYKWKIDLKSEYEKQVDKTKD--KNKMYDNI--DEVKN- 700  
Db 1523 NGSPAIFKPCSAISCRSYKTWINTKKDEKQKQKLYNKEIKONKSNYDNIYDKFVKNL 1582  
Qy 701 ----KEANVYLKEK-----SKECKOVNFDDKIFNES--PNEYEDMC-----735  
Db 1583 CTDYKSVDSFLKLLKEGPCCKNNTKDSKIDFKDTEETFRNAEYCDPCPVFGVICNNGDCS 1642  
Qy 736 ----KKCD--EIKYLNBI-----KYP-----750  
Db 1643 NSTEKKCDAQEFKTYDVKNKENPKNVNLVSDKTAKYPGLDNGVCENSSIFEGIRE 1702  
Qy 751 -----TKKHDIVD-----IDTF---SDTFGGTPIISINANI 778  
Db 1703 KWSGCGYCGLDICPNKNTGDIHDKQNAPIRVLPKRWIENFLKDHKKIKDILSICIN--- 1759  
Qy 779 NEOQSGKDTNMTNGSE-----TSDSPVSH-----EPE 805  
Db 1760 NENRNICDVCCKNCECIDKWIEMKEMKIVRDYVYKQYNVADSVYVYRFLGLOPQ 1819  
Qy 806 SDAAINVEKLSGDESSSETRGILDINDPSVNNVNEVHDASNTQGSVSNSTSDITNGHSES 865  
Db 1820 ND-----LEKVGK-----DVNDLRDLEELSECTNTVSTENRCKRKKQVV---ES 1860  
Qy 866 SLARTTNAQDIKIGRSGNEQSDNQSNSHSDNSGSLTIGQVPSD---NTQNTYDSQN-P 922  
Db 1861 LLNKLKN-----EIRHCKNERDDSGKES-----CKTLPE-PTDDPQTDSDHTDIP 1908  
Qy 923 HRDTPNALASIPSDDKINEIEGFDSSRDSENGRGTTSNTHDVRRTNIVSERRVNSH-DF 981  
Db 1909 PGDVAPTFCNVPA-----NPGCD-KSATNVNVNTEVAKEMHEEAHKOM 1950  
Qy 982 IRNGM--ANNNAHHQYITQIENNGIIRGOESAGSNVYKDNPKRNSFSENHKKVIOE 1039  
Db 1951 LERSVKVESKVDSTVESVLRADASKGYEGHGNPDLDKHNMCNITKEHTNYQKRGYN 2010  
Qy 1040 Y-----NSRDTRV-----RBEILKL-----SKQCNKNEYSNEYCTYSERN 1077  
Db 2011 YRGCTKGNGKOTRFVIGTIWKEDEKDETIVKLLPFRHRMCTS--NLEYLLHVNK-- 2066  
Qy 1078 SSPGPCSREERKLL 1091  
Db 2067 ---GPLLKVPDKI 2077

## RESULT 33

ABB07655

ID ABB07655 standard; protein; 302 AA.

XX ABB07655;

AC ABB07655;

XX 20-MAY-2002 (first entry)

XX P. falciparum EBA-175 protein Pff1 domain sequence.

XX

KW Duffy binding-like domain; DBL domain; parasite; malaria; DBP protein;  
KW protozoacide; vaccine; EBA-175 protein.  
XX Plasmodium falciparum.  
OS WO200212292-A2.  
XX 14-FEB-2002.  
PD 03-AUG-2001; 2001WO-EP009023.  
PF 07-AUG-2000; 2000GB-00019375.  
XX (ITGE-) INT CENT GENETIC ENG & BIOTECHNOLOGY.  
PA Chitnis C, Pandey K, Singh S, Yazdani SS, Pattnaik P;  
PI WPI; 2002-227139/28.  
XX Producing polypeptide with Duffy binding-like domain, by expressing  
PT polypeptide in bacterium/yeast, extracting and denaturing it, refolding  
PT polypeptide in presence of arginine and urea, and optionally recovery.  
XX Disclosure; Page 43-44; 47pp; English.  
PS The invention relates to a method for producing a polypeptide (I)  
CC comprising a Duffy binding-like (DBL) domain. The method involves  
CC expressing (I) in a bacterium, or as a non-secreted polypeptide in a  
CC yeast, extracting the expressed polypeptide from the bacterium or yeast  
CC and denaturing the polypeptide, refolding the extracted polypeptide in  
CC the presence of arginine and urea, and optionally recovering the refolded  
CC polypeptide. The method is useful for producing a polypeptide comprising  
CC DBL. (I) is useful for identifying a substance that modulates the  
CC interaction between the polypeptide and a host cell receptor involved in  
CC the entry of a parasite into a host cell. The substance identified is  
CC useful in the manufacture of a medicament for treating or preventing  
CC malaria. A pharmaceutical composition or a vaccine composition obtained  
CC by formulating the refolded polypeptide is useful for treating or  
CC preventing malaria. (I) is useful as vaccine to prevent malaria or  
CC infection by P. falciparum or P. vivax. The present sequence represents  
CC the P. falciparum EBA-175 protein Pff1 domain sequence  
XX  
SQ Sequence 302 AA;  
Query Match 8.4%; Score 546.5; DB 5; Length 302;  
Best Local Similarity 37.7%; Pred. No. 6.5e-26;  
Matches 122; Conservative 50; Mismatches 119; Indels 33; Gaps 9;  
Qy 121 NTNSNPFANTSEISIGKONKQYTFIOKRTFLPACGIRKRSIKWICRENSKIT-VCVPDR 179  
Db 3 NTSSNN-----EV-----LGNCREKRGKMGKWDCKKQDRSNYVCIPDR 40  
Qy 180 KIQLCVANFLNRLTMEKPKBIFLISVNTAEKLLYNKNEGKDPISFCNELRNSFSDFS 239  
Db 41 RIQLCVINLSIKTYTKETKQDFEASKESQLLLKQDNKYNKSFCDNLKNSFLDYGH 100  
Qy 240 SPTGDDMDFGGNTDRVKGYNITKFSDYKKEKVKELNNIKKEWKEKKNKANLNMHMIWNH 299  
Db 101 LAMGDMDFGGYSTKAENKIQEVFKGAHGEISEHKKIKNFRKEWNEFREKLWEAMLSEHK 160  
Qy 300 GNISKECALIPAEPOINLWIKENENFLMEKKRLFLNIDKCVENKCYACGGRLPC 359  
Db 161 NNIN-NCKNIPOEELQITQWIKEWGHGFEFLERDNRSLPKSKCNNTLYEACEKCIDPC 219  
Qy 360 SSYTSFMKSKTKQMEVLTNLYKKNKSGVDKNFLNDFK-KNNKND--LDDFFKN-EKEY 415  
Db 220 MKYRDWIRSKFEWHTLSKEYETQK--VPKENAENYLIKISEKNDAKVSLLNNDCAEY 277  
Qy 416 DDLCDCRYTATIIKSFINGPAKND 439  
Db 278 SKYCDCKHTTILVKSVLNG---ND 298



QY 932 ----SLPSDDKIN-ELIEGPDSS---RDSNGRGDTTNTHTDVRTNIVSRRV--SHDF 981  
 Db 2626 VDNHTPTTTRHRNVEEKPFTMSIHDRLYSGE-EYSYNVNVANDPISARNGYGGIDL 2684  
 QY 982 IRNGMANN--NAHQYITQIENN--GIIRQERSAGNSV--NYKDNPKRSN---FSGEND 1032  
 Db 2685 INDSLNSKVVDIYDELLKRKENELFGTNHTTKNTSTNSVAKNTNTDPIHNQLNLFHTWLD 2744  
 QY 1033 HKKNIOEYNSRDTRKVRREEII-KLSKO-NKCNN 1063  
 Db 2745 RHRDMCE--KWDNTNNKKEILLKLKEWNNKDN 2775

## RESULT 35

AE822178  
 ID AEB22178 standard; protein; 3346 AA.

AC AEB22178;

DT 22-SEP-2005 (first entry)

DE Plasmodium falciparum VAR6.

KW antimalarial; vaccine; protein purification; pharmaceutical; diagnostic;  
 KW plasmodium falciparum infection; infection; VAR6.

OS Plasmodium falciparum.

XX W02005063804-A1.

XX 14-JUL-2005.

XX 28-DEC-2004; 2004WO-DK000922.

XX 30-DEC-2003; 2003DK-00001954.

XX (UYKO-) UNIV KOENHAVNS PANUM.

XX Jensen ATR, Hvid L, Jorgensen L, Lavsten T, Magistrado P;

XX Nielsen MA, Salanti A, Staalso T, Theander TG;

XX WPI; 2005-522559/53.

XX N-PSDB; AEB22177.

PT New isolated VAR4, VAR5, and VAR6 polypeptides isolated from Plasmodium  
 PT falciparum, useful as vaccines for treating or preventing severe malaria.

PS Claim 1; SEQ ID NO 6; 193pp; English.

XX The invention describes an isolated polypeptide comprising an amino acid  
 CC sequence selected from at least one of SEQ ID NOS 2, 4, and 6 (3553, 2994  
 CC and 3346 amino acids) for use as a medicament. The polypeptides, nucleic  
 CC acids and vectors are useful as vaccines for preventing and/or treating  
 CC malaria. The malaria is preferably severe malaria caused by Plasmodium  
 CC falciparum. An in vitro diagnostic method is useful for detecting anti-  
 CC malarial antibodies. The polypeptides are useful for generating a vaccine  
 CC against severe malaria, where the vaccine comprises antibodies generated  
 CC against the polypeptide which have cross-reactivity to parasites causing  
 CC severe malaria and the ability to inhibit adhesion to endothelial cells.  
 CC The polypeptides are also useful in screening for inhibitor molecules  
 CC capable of inhibiting binding of any of the polypeptides to a receptor  
 CC expressed on endothelial cells. This is the amino acid sequence of  
 CC Plasmodium falciparum Var6 that is transcriptionally upregulated in  
 CC Plasmodium following increased antibody recognition.

XX Sequence 3346 AA;

Query Match 8.4%; Score 543.5; DB 9; Length 3346;

Best Local Similarity 20.8%; Pred. No. 2.8e-24;

Matches 300; Conservative 193; Mismatches 545; Indels 403; Gaps 64;

QY 31 LYNQDESSD---ISRVNSP-----ELNNHKTNIYSDYEDVNNKL 69

Db 730 LNDDEDEDDDETPRAHNPVCDKNDNSOPTKTVSYIARQHRRAQAQMTKNSVVDGNNKL 789  
 QY 70 INSFVENKSVKKRSLSFIN--NKTYSYDIIIPPSYSYRNDKFNLSLENEDNSNTNSNN 126  
 Db 790 EGDIPK-----VTFRGGVGKNGIDICKIDKTSND-----SRGPTDPCGKGG 835  
 QY 127 FANTSISISGKD-----NKQYTFI---OKRTHLPACGIIKRKSIKIWI CHRENSSEKIVCV 176  
 Db 836 ----DRFKIGTDMQGDSPVNPQYRGIYMPRRQH-----FCTSNLEKLDVSR 878  
 QY 177 POKIQLCVANFLNSRLETMEKFEFLISVNTAEKLLYNK--NEGDPGIFCNELNSF 234  
 Db 879 VTRN-----GNASNSLLGDV-----LLAAKYEAERTKNHYVSKKEHSEACRAVYSF 926  
 QY 235 SDFRSSFIGDDM---DFG-----GNTDRV-----KGINTKFSDDYK--EKNVEKLNLIK 279  
 Db 927 ADLGDIIIRGKMWDKNHGEKKTQENLERIFAKIKEQLLNSSIKDKYKDDKATPKYKQLR 986  
 QY 280 KEWWEKXKANLWNHMIWNHK-GNISKECAIIPABE--PQINLMIKEWNEFLMEKKRLFL 336  
 Db 987 EDWWEANRSQVWEAMQCPKNGTFFCKSDHTPLHDYIPQRLRWMTAEWYCKEQSRLYG 1046  
 QY 337 NIKDKCVENKKYEACFGG-----CRLPCSSYTSFMKSKTKQMEVLTNLV-----KKX 383  
 Db 1047 ELVETCGKMHGKCKQGNHCVTCRPAKEYKFFINTWQPMQMEQKYSQLYEAKKY 1106  
 QY 384 NSGVDXN-----NFLNDLFKNNKNDLDD-----407  
 Db 1107 NDSSRKDTTKDDYVLQFLNKLTLQNGKNKYDTAEGYVHQEAHISDCQKQTOPCKKRG 1166  
 QY 408 -----PPKNEKEYDDLDCD--RYTATIIKGFNLGPAKNVDVDSQ-----IN 447  
 Db 1167 EIPSSDTETDNNYAFRPPQPHDHDVECECNTRQTKVRKK-----KKKVDACEMAKTLH 1220  
 QY 448 VND--LRGFGCNYSNNE--KSWNCTGTFNKNPGCTCEPPRQTCGLGHTYLLHRGHE- 501  
 Db 1221 NNDGTIRIGQCKRCKDEGNAEYPKWDCNSQIHHTHNGACMPPRKQLCV---YFANPSQI 1277  
 QY 502 --EDYKEHLLGASIVBAQLLKVK---YKEKD--ENALCSIIQN-----SYAD 541  
 Db 1278 GSINKQDNLKAFIISAAAEFRSQYKSKNGENLQTLQDGTIPDPFKRMFYTYGD 1337  
 QY 542 LADIIKGSIIIDY-YGKQMEENLKNVKNQKGR---NEESLKIIFREKWDENKENVKVM 597  
 Db 1338 YRDLFGTIDISGLGEGTALAKQINILFPNGVRKIPNEKT---REKWWTDHGPEIWKGM 1393  
 QY 598 SAVLKN-----KET--CKDY-----DKFOKIPQFLRWFKWEGDDPCEKR 634  
 Db 1394 LCAITNGLSEKTKIIFDDYSHDKVQSKNGNPSLEDAKQPFPRFWTEWSEDFCRER 1453  
 QY 635 KEKIYSFEKVECKE--KDCDB---NTCKNKSEYKQWIDLKSEYKQVDKYTKDKNK 689  
 Db 1454 KKK---EEVERDKDEYGECEKEKNGKCVTACKAYKEYITNKKEEYDSQKGFDEKTE 1510  
 QY 690 KMYNDIDEVKNKEANYLKEKSKCKDVNFPDDKI FNESPNEDMCKK--CDI IKYLNEI 747  
 Db 1511 KK-QGYEDYSEKQASEYLKEK-----CTKSSCNMKWKVTEI 1545  
 QY 748 -KYPKTKHDIYDITDSTDFG-----DGTPIISINANINQOSCKDTNTNCSSTSDSPVS 801  
 Db 1546 SNTWTPHKTYY-----TENLGIKCECPPSPCTIIVDGLLSPQ-----NSSYANGCKWKYG 1596  
 QY 802 HEPESDAAINVEKLSGDESSSETRGILDINDPSVTNNVNEVHDASNTQGSVNSTDITNG 861  
 Db 1597 KMSQGTEDWDCSKKSGEGEGNEDGVVVCIPPRRRRLYVKNLQDLTGESLV---DLRKA 1652  
 QY 862 HBESSLNRTNA-ODIKIGRSGNEQSDNOENSHSS-----DNGSLTIQGVPSDNTQNT 916  
 Db 1653 FIKCAAIETFFAWHEFKEREKEKNEQDVQYKSVLENLQKLNKGIDIDFEKQRMF 1712  
 QY 917 YDSQNPHRDTPNALASLPDSDDKINEIEGFPDSSRSDENGRGDTTNTHTDVRTNIVSRRV 976



QY 652 -----DCDENTCKNCKSEYKWKIDLKSEYKQVDKYTKDKNKKMYDNIDE 697  
 Db 2114 ACNEMLPKNDGTVPDLLEKPSCAKPSYRKWIESKGEKFEKQKAYEQQDKCVNGS--- 2170  
 QY 698 VKNKEANVYLK-----EKSK-----CKDVNF-----DDKIFNESPN-----EYED 733  
 Db 2171 --NKHDNGFCETLTSTSKAKDLTKLGPCKPNVVEGKTIIDDDKTFKHTKDCDPLKFSV 2228  
 QY 734 MCKK--CDEIKYLNKIKYPTKHDIYDI-----DTFSDTFDGTGTPISINA 776  
 Db 2229 NCKKDCDCKSGTD--CRNKNSIDATDIENGVDSTVLEMRVSADSKSGFGDGLGNACRG 2286  
 QY 777 -----NINEQSQSKO----- 786  
 Db 2287 AGFEGIRKDEWKCRNVGVVCKPENVNGEAGKHIIQIRALVWRKRVVEYFFDYNNKIKH 2346  
 QY 787 --TSNTGNSETS-----DSPVSEHPESDAAINVEKLSGDESSSETRGILD--- 829  
 Db 2347 KISHRIKNGEISPCIKNCVKEKWDQKKEWKEITERFKQYKNDNSDDDNVRSFLETLP 2406  
 QY 830 -INDPSVTNNVNEVHDASNTQG--SVSNTSDITNGHSESSLNRTTNAQDIKIGRSGNEQSD 887  
 Db 2407 QITDANAKNKVILKSKFGNSCGCSASANEQNKNGEYKDAIDCMLKKLKDKGEC--EKKH 2464  
 QY 888 NOENSSHSDNSGLTI-----GOVPSD 911  
 Db 2465 HQSDTECSDTPOQTLEDETLDDIETEBAKNMMPKICENVLTKAQBSDEGCVPAPEN 2524  
 QY 912 N-----TONTYDSQNP 923  
 Db 2525 SEBPAATDSKETPEQTPVLKPBEEAVPEPPPPQEKAPAPIPOQPPTPTQLLDNPH 2584  
 QY 924 -----RDTPNALASL-----PSDD-----KINE----- 941  
 Db 2585 VLTALVTSLAWSVGIGFATFTFYLLKKTSSVGNLFQILQIPKSDYDIPTKLSPNRYI 2644  
 QY 942 -----IEGFDSSRDS--ENGRGDTSTNTHVRRTNIVSERVNSHDFIRNG 985  
 Db 2645 PYTSGYGRKYRYILEG--DSGTDGTYDHYSDITSSSEYEEMDI-----NDIYVPG 2695  
 QY 986 MANNNAHQYITOIE-----NNGIIRGOESA-GN--SVNYKDNPKRNSFSEND--- 1032  
 Db 2696 SP-----KYTLIEVLEPSGNNTTASGNNTTASGNNTTASGNNTSDQNDIQNDGIP 2749  
 QY 1033 -HKNTIQEYNSRDTKRVREBIIKLSKQNCNNEYSMEYCTYSDERNSSPGCSRE-ERK 1090  
 Db 2750 SSKITDNEWN-----QLKDSFISQYLQSEPNTEPNM--LGYNVDNNHTPTTSHHVEEK 2802  
 QY 1091 LCCQISDYCLKYFNFSIYNYCNKSEIKSPYKCFKSEGOSSIPYPAAGILVIVILL 1150  
 Db 2803 FIMSIIHDR-----NLFSGEYN-----YDMFNS--GNPFI----- 2830  
 QY 1151 SSASRMGKSNEEYDIGESNTEATFEENNYLNKLSRIFNQVQETNIDSYSEYNYNEKN 1208  
 Db 2831 ----NISDSTNSMDSLTSSNHSFNDKNLYSGIDLINDALSNGHIDIYDMLKRKEN 2884  
 RESULT 37  
 ID AAY77905  
 XX AAY77905 standard; protein; 3060 AA.  
 AC AAY77905;  
 XX AAY77905;  
 DT 13-JUN-2000 (first entry)  
 XX Plasmodium var-7 polypeptide.  
 DE DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;  
 KW DABP; Sialic Acid Binding Protein; SABB; malaria; vaccine; immunisation;  
 KW protozoacide; var-7.  
 XX Plasmodium sp.  
 OS

XX US5993827-A.  
 PN 30-NOV-1999.  
 XX 07-JUN-1995; 95US-00487826.  
 XX 10-SEP-1993; 93US-00119677.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX Sim KL, Chitnis C, Peterson DS, Su X, Wellem TE, Miller LH;  
 DR WPI; 2000-194198/17.  
 DR N-PSDB; AA298287.  
 XX Isolated protein binding domains from Plasmodium vivax and Plasmodium  
 PT falciparum erythrocyte binding proteins useful for vaccinating against  
 PT malaria.  
 XX Disclosure; Col 109-124; 93pp; English.  
 PS The invention relates to ebl-1 polypeptides that are encoded by the DBL  
 XX (Duffy-binding like) gene family. The ebl-1 proteins are substantially  
 CC identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid  
 CC Binding Protein (SABP), which are soluble proteins that appear in the  
 CC culture supernatant after erythrocytes infected with malaria release  
 CC merozoites. Immunohistochemical studies indicate that DABP and SABP are the  
 CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy  
 CC and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be  
 CC used to vaccinate against malaria, especially caused by P. falciparum.  
 CC Immunization with the polypeptide provides effective protection against  
 CC malaria. The present sequence represents the var-7 polypeptide  
 XX  
 SQ Sequence 3060 AA;  
 Query Match 7.7%; Score 498.5; DB 3; Length 3060;  
 Best Local Similarity 18.9%; Pred. No. 1.7e-21;  
 Matches 317; Conservative 180; Mismatches 520; Indels 661; Gaps 67;  
 QY 90 NKTYSYDIIPPSYVRNDFNSLSENEGNTNSNNTSE-ISIGKDNKQVTFIQKR 148  
 Db 1309 NKEYTFQPPPEVATACDINRSQTEPKKEENVESACKIVKILEGKNG-----R 1360  
 QY 149 THLFAGCIKRSIKWICRENS--KITCVDPKIQLCVANFLN----SRLETWEKPEI 202  
 Db 1361 TTVGECNPKESYPMDCNKNIDISHDGCACMPRRQKLCLYYIAHESQTEIKTDNLDKA 1420  
 QY 203 FLISV-----NTEAKLLYNKNEGKDPISFCNELRNSFSDFRSFIGDDMD 247  
 Db 1421 FIKTAAAEFTLSQWYKSKNDSEAKIL---DRGLIPSQFLRSMYTFGDYR-----D 1469  
 QY 248 FGQNTDRVGYNTKPSDYKQVKEKNVEKLNMI-----KKEWKEKNKANLWNH 294  
 Db 1470 ICLNTD-----ISKQNDVAKAK--DKICKFSGKSGKSPGSLRSGEWWKTNPEIKWGM 1522  
 QY 295 IV-----NHKGNISKECAIIPAEPEQINLWIKWENENFL 328  
 Db 1523 LCALTKYVTDTDNKRKIKNDYSYDKVQNGNPSLEEF---AAKQFLRWMIENGSEFC 1579  
 QY 329 MEKKRLFLNKKCVENKKEACFGG---CRLPCSSYTSFMKSKTKQMEVLTN----- 378  
 Db 1580 AERQKKEIKKACNEINSTQCNDAKHCNQACRAYQEVYENKKEKFSQTNFVLKAN 1639  
 QY 379 -----LYKKKNSGVDKNNFLNDFKQKNKDL-----DFFKNEKEYD 416  
 Db 1640 VQPDPEYKGYKDGQVPIQNGEYLLQKCDNNKSCMDGNVLSVPKPEKPKYAHKYP 1699  
 QY 417 DLCDCRYTATIIKSFUNGPAKND-----VDIASQINLV---NDLRGF---GCNYSKN 461  
 Db 1700 EKDCD-YQKGKHPVSIPIPPPPVQPEAPTPTVDVCSIVKTLFKDTNNTFSDACGLKYGKT 1758  
 QY 462 NEKSNWCTGTFTNKPPGT-----CEPPRQTCLGR----- 492



1067 EKHDGKDNPKYTGDESKPAYK - KLRADWBEANRHQVWRAMKCATKGI - --CPGMPV 1121  
312 EE--POINLWIKENWENFLMEKRLFLNIIKDKCUE--NKYIEAC-----FGGRILPCSS 361  
1122 DDYIPQRLRWMTWEAWYCKAQSOEYDKLKKI CADCMKSGDKGCKTQGVDCGRKCAADK 1181  
362 YTSFMKSKTKQMEVLNLY--KKNKSGVDKN-----NFLNDLFK--- 398  
1182 YKEIEKWNQWRIKIDKYNLLYLQAKTTSTNPGRTVLGDDDDPYOQMVDFLPIHKASI 1241  
399 ----- 398  
1242 AARVLVKRAAGSPTEIAAAAPITPYSTAAGYIHQEIYGGCQEQQTQCEKKHGATSTTT 1301  
399 KNNKNDLDDFFKEKEVDLDCRYTATIIKSLFNGPAKNDVDIASQINVD--LRGFG- 455  
1302 KENKE--YTQKPPPEYATACD-----INRSOTEPEKKEENVESACKIVEKILEGKNG 1354  
456 -----CNYKSNNEKSNCTGTFTNKPFGTCEPPROTLCIGRTYLLHRGHEEDYK--EH 507  
1355 RTTVGECNPK-ESYPDWDCQKNIDISHGACMPRRQKJCL--YYIAHESQTENIKTDNN 1411  
508 LLGASIEYA---QLLKYKEDENALCSIION-----SYADLADIIGSDI 551  
1412 LKDAFIKTAATAETFLSQYKSKNDSEAKILDRGLPSQFLRSMYTFGDRDILCLNTDI 1471  
552 IKDYIG-KQMEENLNK-VNKDKRNEESLKIPEKRWENKENVKWMVAVL-----K 602  
1472 SKQNDVAKADKIGKIFSKDGSKSPSGLS--RQEWKKTNGPEIWKMLCALTKYVTDTD 1529  
603 NKETCKD---YDK-----POKIPQFLWPKWEGDDFCERKEKISYFSPFKV 646  
1530 NKIKINDYDKNVQSONGNSPLEEPAAPQFLRMWIEGEEFCAERQKK--ENIIKD 1586  
647 ECKK----KDCDE--NTCKNKSEYKWKIDLKSEYEKQVDKYT-----KDKNKQMYD 693  
1587 ACNEINSTQCNDKAHRCNAQRAYQEYVENKKKEFSGQNTNVLKANVQPPQPEYKGYE 1646  
694 NIDEVKNKEANVLKESKECKDVNFDKIFNESPNE-----YEDMKCKDEIKYL 744  
1647 YKGVQPIQGNVLLQCDNNKSCMDGNVLSVPKEKPFKYAHKYPEKC-DCYQGHV 1705  
745 NETKYPK-----TKHDIYDIDTFSDT-----FGDGTPIINANINEQ 781  
1706 PSIPPPPPVQPOEAPTVTVDCSIVKTLFKDTNNFSDACGLKYKGTAPSSWKCPSPD 1765  
782 QSGKDTNTNGSETSDSPV-----SHEPESD----- 807  
1766 KSGAG-ATTCKSGSDSGSICIPRRRLYYVKLQEWATALPQGEGAAPSHSRADDLRNF 1824  
808 ----AAINV-----EKLSDGESSETRGIL-----DINDPSVTNNVNEVHDASNTQ 849  
1825 IQSAAIETFLWDRYKEEKPGDGSQOALSQLTSTYSDEEDPP-----DKLLQN 1875  
850 GSV-----SNTSDI--TNGHSESSLNRTTNAQDIKIGRSGNEQ 885  
1876 GKIPDPFLRMFTLGDYRDLVHGNTSOGTNG-----SNNNVILVLEASNGKE 1926  
886 SDN--QENSSHSDNSGSLATI-----QQVPE---DNTQN-----TYDSQNPHRD 926  
1927 DMQKIQEKIEQILPKNGGTLVPKSSAQTDPKWNEHAESIWKGIMICALTYTEKNPDTSA 1986  
927 PNALASLPDDKINEIGFPOSSRDSSENGRGDTTSNTHDVRRTNIVSERRVNSHDFTRNG 986  
1987 RGDENKIEKDEV--YEKFFGADKHGTASTTGT----- 2020  
987 ANNAHQQYIQTENNIGIRGOEESAGNSVNYKDNPKRNSFSSNDHKNIQEVNSRDTK 1046  
2021 -----YKQDYKVKLEDTSGAKTPSASDTPLLSDFLVLRPPYFRVLEWGNQFCK 2072  
1047 RVREBIKLSKQNK-----CNNEYSMEYCTYSD-ERNSSPGP 1082  
2073 KRKHQAQIKHECKVEBNGGSRGGITROYSGDGEACNEMLPKNDGTVFDLEKPSCAKP 2132

QY 1083 CSR-----BERKKLCCQISDYCLKYENFYFYSIEYNYCIKSEIKSP 1123  
Db 2133 CSSYRKWIESKGKEFEKQKAYEQKDKCVNGSNKHDNGFCETLTTSKAKDF 2185

RESULT 39  
AAW22482

ID AAW22482 standard; protein; 2710 AA.

XX AC AAW22482;

XX DT 07-OCT-1997 (first entry)

XX DE Plasmodium Proj3.

XX KW DBL gene family; SABB; sialic acid binding protein; vaccine; therapy;  
XX KM Duffy binding like gene; Duffy antigen binding protein; erythrocyte;  
XX KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;  
XX KW Plasmodium.

XX OS Plasmodium falciparum.

XX PN WO9640766-A2.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US009508.

XX PR 07-JUN-1995; 95US-00487826.

XX PA (USSH ) US DEPT HEALTH &amp; HUMAN SERVICES.

XX PI Sim Kl, Chitnis C, Miller LH, Peterson DS, Su X, Wellems TE;

XX DR WPI; 1997-052231/05.

XX DR N-PSDB; AAT72897.

XX PT New malaria vaccines - contains cysteine-rich DBL family protein binding  
XX PS domains homologous domains of the Duffy and sialic acid binding proteins.  
XX PS Disclosure; Page 50-56; 96pp; English.

XX CC This sequence represents Proj3 of Plasmodium. Proj3 belongs to the Duffy  
XX CC binding like (DBL) family of genes which have homology to the Duffy  
XX CC antigen binding protein (DABP) and sialic acid binding protein (SABP)  
XX CC conserved regions (see AAT72889 and AAT72888 respectively). The var  
XX CC family of genes modulate cytoadherence and antigenic variation of  
XX CC plasmodium infected erythrocytes. SABP and the Duffy antigen binding  
XX CC protein (DABP) are soluble proteins that appear in the culture  
XX CC supernatant after infected erythrocytes release merozoites. DABP and SABP  
XX CC mediate the binding of merozoites and schizonts to the erythrocyte  
XX CC surface. These proteins are necessary for erythrocyte invasion by the  
XX CC parasite. This sequence can be used in the compositions of the invention.  
XX CC The compositions are for the treatment and prevention of malaria, and  
XX CC comprise either a nucleotide sequence or encoded polypeptide of the var-  
XX CC 1, var-2, var-3 or var-7 genes of the DBL gene family, a family of genes  
XX CC having homology with conserved regions of DABP and SABP. The compositions  
XX CC are used for the treatment and prevention of malaria. They are also used  
XX CC in the preparation of vaccines for inducing a protective immune response  
XX CC in a mammal to Plasmodium merozoites (especially Plasmodium falciparum or  
XX CC Plasmodium vivax)

XX SQ Sequence 2710 AA;

## Query Match

7.6%; Score 495.5; DB 2; Length 2710;

Best Local Similarity 19.5%; Pred. No. 2.2e-21;

Matches 268; Conservative 170; Mismatches 482; Indels 453; Gaps 57;

QY 112 LSENNEDNGNTSNFNTANTSEISIGNDKNQYTFIQKRTLFLACGIK-RKSIKWICRNSE 170

Db 912 INENYSDNRSGNSGGPCT-----GKGDH-----GGVRMRIGTENSIEGKK 953

Qy	171	KIT----	VCVPRKIQLCVAFNPLSRLETM----	EKFKEIFLIVNTEAK----	213		
Db	954	QTSYKNVFLPRPREHMCTSNLENDVGSVTKNDKASHSLGDLQOLAQAAEIIKRYKD	1013				
Qy	214	----	LYNKNEGKDPISFCNELANSFSDFRSSFTIGDDM----	DFGNTDTRVKGYI----	259		
Db	1014	QNNIQLTDP1QOQDOEAMCAVRYSPADLGD1IRGRDWDDEKSDTMEFLTVFKNIK	1073				
Qy	260	----	NTKFSFYKQKVEKLN1IKGBWWEKKNLANH1VNHKNISKECALIPA	311			
Db	1074	EKHGIDKNDPKYTGDSEKPAYK--KLRADWWEANRHQVWRAMKCATKGI1--	CPGMPV	1128			
Qy	312	EE--	POINTLIKWENFLMEKRLFLN1KDKCVB--	NKKYEAC----	FGGCLRPCSS	361	
Db	1129	DDIYPOBLRWTEWAEWYCKAQOQYDKLKK1CADCMKSGDKCTQGDVDCGCKCAACDK	1188				
Qy	362	YTSFMKSKTKQMEVLTWLY----	KKNSGVDXN----	NFLNDLFK----	398		
Db	1189	YKEIEKWNQWRKISKYNLLV1QAKTTSTNPGRTVLGDDDPYQOMVDFLPIHKASI	1248				
Qy	399	----	----	----	398		
Db	1249	AARVLVKRAAGSPTEIAAAAAPIPTYSTAAGYIHOE1GYGCGQEQTOQCEKXKHGATSTTT	1308				
Qy	399	KNNKNDLDDFFKEKEYDLDCCRYTATI1KSFPLNGPAKNDVDIASQINVD--	LRGFG--	455			
Db	1309	KENKB--	YTPKQPPPEYATADC-----	INRSOTEPEPKCEENVSACKIVEXILEGKNG	1361		
Qy	456	----	CNYKSNNEKSNWCTGTFTNKPGPCEPPROTCLGRTYLLHHRGHEBDYK--	EH	507		
Db	1362	RTTVGECNPK--	ESYPDWDCNN1D1SHDGACMPRRQKCL--	YYIAHESQTENIKTDN	1418		
Qy	508	LLGAS1YEA--	QLLKYYKKEBDENALCS1ION-----	SYADLADI1KGSDI	551		
Db	1419	LKDAFIKTAABATFLSMQYKSKNDSEAKILDRGLIPSOFLRSMYVTFDGYRDI	CLNTDI	1478			
Qy	552	IKDYYG--	KMEENLNK--	VNKKRNEESLKI	PREKXWDENKENVWKVMSAVL-----	K	602
Db	1479	SKQNDVAKAKD1K1GRFPFSKDGSKSPGLS--	RQEWKTINGPE1WKGMLCALTKYVTD	1536			
Qy	603	NKETCKD----	YDK-----	PQK1PQFLRPFKEWDDPCEKREK1VSPESFKV	646		
Db	1537	NKKK1KNDYSDYKVNQSONGNPSLEEPAKPQFLRM1EWEEFCAERQK--	ENI1KD	1593			
Qy	647	ECKK----	KDCDE--	NTCKNKCSEYKKW1DLKXSEYKQVDKYT-----	KDNKKMYD	693	
Db	1594	ACNEINSTQCNDAKHRCNOACRAYQEVYENKKEKFSGQTNNFVLKANVQOPQPEYKGYE	1653				
Qy	694	NIDEVKNKEANVYLKEKSEKCDVNPDDK1FNSPNE-----	YEDMCKKCEDEIKYL	744			
Db	1654	YKDGVOPIOGNEBYLLOKCDNNKSCMDGNVLSVPKEKPFKGYAHKYPEK--	DCYQGHV	1712			
Qy	745	NE1KYPK-----	TKHDIYDIDTFSDT-----	FGDGTPI1SINANINEQ	761		
Db	1713	PS1PPPPPPVQOPEAPTVTVDVCS1VKTLFKDNTNFSADGLKYGKTAPSSWKC1PSDT	1772				
Qy	782	QSGKDTN1GNTS1SDSPV-----	SHEPESD-----	807			
Db	1773	KSGAG--	ATTCKSGSDSGS1CIPRRRLYVYK1QEWATALPQEGEAGAPSHSRADDLR	1831			
Qy	808	----	AAINV-----	EKLGDGESSSTRGIL-----	DINDPSVTNNVNEVHDASNTQ	849	
Db	1832	IQSAAIETFLWDRYKEEKKQGDGSGQALSQLTSTVSDDEEDPP-----	DKLLQN	1882			
Qy	850	G5V-----	SWTSD1--	TNGHSSLSNRTPPTTAAQDIK1GRSGNEQ	885		
Db	1883	GK1PPDFLR1MFYTLG1DYR1LVHGNTSDSGTNG-----	SNNNN1VLEASGNKE	1933			
Qy	886	SDN--	QBNSSHSDNSGLTI-----	GQVPSE--	DNTQN-----	TYDSONPHRDT	926
Db	1934	DMQK1KE1BQ1I1PKNGGTF1VPKSSQ1TDPKWNNEHAE1W1KGM1CAL1Y1TBKNPD1SA	1993				
Qy	927	PNALAS1PDSDDK1NE1EGFDSRSDENGRGD1T1SNTHDVR1T1V1SERRVNSHDF1IRNGM	986				

[illegible]

100

```
Db 912 INENYSNDRSGNGPCT-----GKGDH-----GVNRIGTEWSNEGKK 953
Qy 171 KIT---VCVPDRKIQLCVANFLNSRLETM---BKFEKIFLISVNTAKL-----213
Db 954 QTSYKXNVFLPPREHMTCSNLENLDVGSVTKNDKASHSLLDGVQLAAKTDAAEIIKRYKD 1013
Qy 214 -----LYNKNEGKPSIFCNELNSPSDFRSSIFGDDM---DFGNTDRVKGYI-----259
Db 1014 QNNIQLTDPITQQKQDOEAMCAVRYSFADLGGIIIRGRDMWDEKSDTDMETRLITVFNKIK 1073
Qy 260 -----NTKFSDDYKKEKVKELNNIKKEWKEKNKANLWNHIVNHKGISKECALIPA 311
Db 1074 EKHDGIGKDNPKYTGDESKEPAYK--KURADWNEANRHQVWRAMKCATKGII---CPGMPV 1128
Qy 312 EE--PQINLWIKENENFLMEKKRLFLNIKDKCWE--NKYEAC-----FGSCRIPCSS 361
Db 1129 DDYIPQRLRWMTWEAWYCKAQOQYDKLKKICADCMKSGDKCTQGDVDCGKCAACDK 1188
Qy 362 YTSFMKSKTQMEVLTNLY-----KKNSGVDKN-----NFLNDLPK---398
Db 1189 YKEIEKWEQWRKISKDKNLLYLQAKTTSTNPGRTVLGDDDPDYQQWVDFLPIHKASI 1248
Qy 399 -----398
Db 1249 AARVLKRAAGSPTEIAAAPITPYSTAAGYIHQEIYGGCQBQTOFCEKKHGATSTTT 1308
Qy 399 KNNKNDLDDFFKNEKEYDDCLDCRYTATIIKSLFNGPAKNDVDIASQINVD--LRGFG- 455
Db 1309 KENKE--YTFKQPPPEVATADC-----INRSQTEEPKKEENVEACKIVEKILEGKNG 1361
Qy 456 -----CNYKSNNEKSNWCTGTFTNKPFGTCEPPRQTLCLGRYTLHLRGHEEDYK--EH 507
Db 1362 RTTVGECNPK--ESYPDWDCNNIDISHDGACMPRPQKCL--YYIAHESQTEENIKTDDN 1418
Qy 508 LIGASIVEA---QLLKYKYEKENALCSIIQN-----SYADLADIIGSDI 551
Db 1419 LKDAFTAAAEFTLSQWYKSKNDSEAKILDRGLIPSQFLRSMYTFGDIYRDICLNTDI 1478
Qy 552 IKDYYG-KMMEENLNK-VNKDKRNEESLKFREKWDENKENVWKMVAVL-----K 602
Db 1479 SKQNDVAKAKDKIGKFFSKGSKSPGLS--ROEWKNTGPEIWKMGCALTKYVTDTD 1536
Qy 603 NKETCKD---YDK-----FQKIPQFLRMFKWGGDDFCBKRKEKIYSPESFKV 646
Db 1537 NKRKIKNDYSYDKVNSQNGNPSLEEPAAKQFLRWMIWGEFECAERQKK---ENIIKD 1593
Qy 647 ECKK----KDCDE--NTCKNKCEYKKWIDLKSEYEKQVDKYT-----KDKNKKMYD 693
Db 1594 ACNEINSTQCNDAKHKRCNOACRAYOBYENVENKKEFSGOTNNFVLKANVQPDPEYKGYE 1653
Qy 694 NIDEVKNKEANVYLKESKECKVDNFDKIFNESPNE-----YEDMCKKDEIKYL 744
Db 1654 YKDGVOPIQNEVLLQKCDNNKCSMDGNVLSVPKPEKPGKYAHKYPEKC--DCYQGHV 1712
Qy 745 NEIKYPK-----TKHDIYDIDTFSDT-----FGDGTPIISANINEQ 781
Db 1713 PSIPPPPPVQPOPEAPTVTVDVCSIVKTLFKDNTNPFSDACGLKYGTAPSSWKCIPSDT 1772
Qy 782 QSGKOTSGNTGSETSDSPV-----SHEPESD-----807
Db 1773 KSGAG-ATTGKSGSDSGSICIPRRRLYYVKLQEWATALPOGEGAAPSHSRADDLRNF 1831
Qy 808 ---AAINV-----EKLSGDESSESTRGIL-----DINDPSVTNNVNEVHDASNTQ 849
Db 1832 IQSAAIETFFLWDRYKEEKPOQDGSQOALSQITSTYSDDDEDP-----DKLLQN 1882
Qy 850 GSV-----SNTSDI--TNHSESSLNRTTNAODIKIGRSGNEQ 885
Db 1883 GKIPDPFLRLMFTYTLGYRDLVHGGNTSDSGNTNG-----SNNNNIVLEASGNKE 1933
Qy 886 SDN--QSNSSSHSDNSGSLTI-----GOVPSE---DNTQN-----TYDSQNPHRDT 926
```

```
Db 1934 DMQKIQEKIEQILPKNGGTPLVPKSSAQTPDKWNEHAESIWKGMICALITYTEKNPD TSA 1993
Qy 927 PNALASLPDSDDKINEIEGFDSSRDSSENGRGDTTNTNTHDVRRTNIVSERRVNSHDFIRNGM 986
Db 1994 RGDENKIEKDDEV--YEKFFGSTADKHGTASTPTGT-----2027
Qy 987 ANNNAAHQVITQIENNGIIRGOEBSAGNSVNYKDNPKRNGFNSSENHKKNIQIYNSRDTK 1046
Db 2028 -----YKTQDYKVKVLEDTSGAKTPSASDTPLLSDFLRPPYFRYLEWQGNFCK 2079
Qy 1047 RVREELIILSKONK-----CNNEYSMEYCTYSD--ERNSSPGP 1082
Db 2080 KKKHKLAIQIKHECKVVEENGSGRRGIGITROYSGDGEACNEMLPKNDGTVPDLEKPCAKP 2139
Qy 1083 CSR-----BERKKLCCQISDYCLKYFNFYISIEYINCISEIKSEI KSPEY 1123
Db 2140 CSSYRKWIESKGKEFEKQEKAYEQKDKCVNGSNKHDNGFCETLTITSSKAKDP 2192
```

Search completed: November 21, 2005, 20:33:45  
Job time : 230 secs